

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 13:45:21 ; Search time 8024 Seconds

(without alignments)
11540.127 Million cell updates/sec

Title: US-10-602-898A-1

Perfect score: 1629

Sequence: 1 cctgcagtagcagctgttg.....cagaagataaaacgctacta 1629

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_in.*

3: gb_env.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pr.*

9: gb_ro.*

10: gb_ete.*

11: gb_ev.*

12: gb_un.*

13: gb_vl.*

14: gb_htg.*

15: gb_pl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1629	100.0	1629	15	ATU12232	U12232 Arabidopsis
2	1536.4	94.3	1619	15	AY045860	AY045860 Arabidopsis
3	1132.4	69.5	1134	6	CQ804428	CQ804428 Sequence
4	1113.8	68.4	1128	15	BT020366	BT020366 Arabidopsis
5	811.8	49.8	125502	15	AT74L20	AL023094 Arabidopsis
6	811.8	49.8	195921	15	ATCHRIV81	AL161585 Arabidopsis
7	661.6	40.6	1526	15	AF170921	AF170921 Pisum sat
8	656.4	40.3	1430	15	NT284820	Z84820 N.tabacum m
9	654.8	40.2	1600	15	AF249743	AF249743 Nicotiana
10	651	40.0	1434	15	NT284821	Z84821 N.tabacum m
11	644.8	39.6	1505	15	NPY09513	Y09513 N.plumbagin
12	642.4	39.4	1540	15	AF414114	AF414114 Solanum t
13	641.2	39.4	1560	15	NTGPRTEET	X98161 N.tabacum m
14	640.8	39.3	1524	15	STGB1GENE	X87837 S.tuberosum
15	640.6	39.3	1611	15	AF145976	AF145976 Pisum sat
16	576.4	35.4	1664	15	OSGPROTBS	X89737 O.eativala mR
17	558.8	34.3	1671	15	ZMU12233	U12233 Zea mays GT
18	558.2	34.3	1688	15	AB090160	AB090160 Triticum

19	551.8	33.9	1470	15	AF033357	AF033357 Avena fat
20	422	25.9	1157	15	BT017412	BT017412 Zea mays
21	385.8	23.7	1722	15	NTU22524	U22524 Nicotiana t
22	344.8	21.2	1777	15	PPA428878	AJ428878 Physcomit
23	284.4	17.5	4090	15	AF533440	AF533440 Pisum sat
24	257.4	15.8	110000	15	AP008209	Continuation (262
25	257.4	15.8	110000	15	AP008209	Continuation (263
26	257.4	15.8	139187	15	AC133930	AC133930 Oryza sat
27	229	14.1	98126	14	AP008019	AP008019 Lotus cor
28	187	11.5	1536	15	PPA512651	AJ512651 Physcomit
29	167.8	10.3	1106	15	AY599846	AY599846 Rhizopus
30	167.6	10.3	1431	15	AF306565	AF306565 Pneumocys
31	151.2	9.3	3259	5	BC076910	BC076910 Xenopus t
32	151	9.3	2260	15	AB072452	AB072452 Fusarium
33	147.8	9.1	3947	15	AY219172	AY219172 Fusarium
34	146.4	9.0	1982	2	AK114926	AK114926 Ciona int
35	146	9.0	1394	2	LFBGTPB	X56757 L. forbesi
36	145.6	8.9	1044	2	DGGBS	X73641 Dictyosteli
37	144.8	8.9	3112	5	BC084263	BC084263 Xenopus l
38	144.6	8.9	2634	15	CPU95139	U95139 Cryphonectr
39	144.4	8.9	1026	2	AY389499	AY389499 Loligo ps
40	144.4	8.9	1619	5	AF277161	AF277161 Ambystoma
41	143.8	8.8	3354	15	AB188197	AB188197 Botryotin
42	143.2	8.8	1655	5	XLKGBETA1	X86969 X.laevius mR
43	142.4	8.7	1611	5	CR761620	CR761620 Xenopus t
44	142.2	8.7	1600	15	AY661671	AY661671 Lentinula
45	140.6	8.6	1730	2	LSGPBS	Z23105 L.stagnalis

ALIGNMENTS

RESULT 1	ATU12232	1629 bp	mRNA	linear	PLN 12-OCT-1994
LOCUS	Arabidopsis thaliana Columbia GTP binding protein beta subunit				
DEFINITION	(AGB1) mRNA, complete cds.				
ACCESSION	U12232				
VERSION	U12232.1				
KEYWORDS	GI:557693				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 (bases 1 to 1629)				
AUTHORS	Weiss, C.A., Garnaat, C.W., Mukai, K., Hu, Y. and Ma, H.				
TITLE	Isolation of cDNAs encoding guanine nucleotide-binding protein beta-subunit homologues from maize (ZGB1) and Arabidopsis (AGB1)				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 91 (20), 9554-9558 (1994)				
PUBMED	7937804				
REFERENCE	2 (bases 1 to 1629)				
AUTHORS	Ma, H.				
TITLE	Direct Submission				
JOURNAL	Submitted (12-JUN-1994) Hong Ma, Delbruck, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724-2212, USA				
FEATURES	Location/Qualifiers				
source	1..1629				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				
	/db_xref="taxon:3702"				
	/clone_lib="lambda YES based cDNA library from above soil organs; Ellledge et al., Proc. Natl. Acad. Sci. USA 88, 1731-1735 (1991)."				
gene	1..1629				
	/gene="AGB1"				
5'UTR	1..207				
	/gene="AGB1"				
CDS	208..1341				
	/gene="AGB1"				
	/codon_start=1				
	/product="GTP binding protein beta subunit"				

/protein_id="AA50445.1"
/db_xref="GI:557694"
/translation="MSVSELKHERHAVATETVNNLRDQLRQRLLQLDQDVARVSAAGQ
RTRVSGATDLVCCRTLQKGYSLDWTPEPNNRIVASQDGRLLVNNALTSOKTHA
IKLPCAMWMTCAFPNGQSVACGGLDSVCSIFSLSSADKDGTPVPSRMLTGRHGYVS
CCQVVPNEADHLITSSGDOTCILMDVFTGLATSVGEPFQSGHTADVLVSISGSNPN
WPIGSCDSTARLMDTRASRAVTRFHGHEGDVNTVFPDGYRFGTSGSDGTCRLYD
IRTHQJUVQYPHGDGNGPVTISAFVSGLLTFAGVYASNNTCYVMDTLLGVLDLIG
LQDQSHRNRIISCLGSLDGSALCTGSDNSNLKIWAFGGHRRVI"
1342..1629
/gene="AGB1"
1629
polyA_site
/gene="AGB1"
/note="9 A residues; alternate polyA sites at 1555, 1586,
and 1606"

ORIGIN

Query Match 100.0%; Score 1629; DB 15; Length 1629;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1629; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCTGACGTAGCAGCGTGTGTGCTTGACTGATCTCTCTCAAGCTTTTTTAATCTCTC 60
Db 1 CCTGAGGTAGCAGCGTGTGTGCTTGACTGATCTCTCTCAAGCTTTTTTAATCTCTC 60
Qy 61 TCTCTTTTCCACGTAATTCCTCCCAAATCCATCTCTTCTAGGTTTCGATCCCTCTCTC 120
Db 61 TCTCTTTTCCACGTAATTCCTCCCAAATCCATCTCTTCTAGGTTTCGATCTCCCTCTCTC 120
Qy 121 AATCATGAACCTCTCTCTAGACCCCAAGTTTCCCTTTTATTTGATCGGCG 180
Db 121 AATCATGAACCTCTCTCTAGACCCCAAGTTTCCCTTTTATTTGATCGGCG 180
Qy 181 ACGGAGAACCTTAAGTCTGATCCCGGAATGTCTGTCTCCGAGCTCAAGAAACGCCACGCC 240
Db 181 ACGGAGAACCTTAAGTCTGATCCCGGAATGTCTGTCTCCGAGCTCAAGAAACGCCACGCC 240
Qy 241 GTCGTACGGAGACCGTTAAATACCTCCGTGACACAGCTTAGACAGACGCTCCAGCTC 300
Db 241 GTCGTACGGAGACCGTTAAATACCTCCGTGACACAGCTTAGACAGACGCTCCAGCTC 300
Qy 301 CTCGATACCGATGTGGCAGGATTTACGGCGGCAAGACGCTACTCGGGTGAGCTTCGGA 360
Db 301 CTCGATACCGATGTGGCAGGATTTACGGCGGCAAGACGCTACTCGGGTGAGCTTCGGA 360
Qy 361 GCAACGGATCTGGTTGTGTGCTACTCTTCAGGGAACACCGGAAAGGTTTATTCATTA 420
Db 361 GCAACGGATCTGGTTGTGTGCTACTCTTCAGGGAACACCGGAAAGGTTTATTCATTA 420
Qy 421 GATTGGACACCGGAGAGAACCGGATTCAGTGCACTCAAGATGGGAGATTAATCGTG 480
Db 421 GATTGGACACCGGAGAGAACCGGATTCAGTGCACTCAAGATGGGAGATTAATCGTG 480
Qy 481 TGGAAATGCTCTAAACGAGTCAGAAACTCATCTATTAACCTCCCTGTGCAATGGGTTATG 540
Db 481 TGGAAATGCTCTAAACGAGTCAGAAACTCATCTATTAACCTCCCTGTGCAATGGGTTATG 540
Qy 541 ACATGTGCTTTCTCTCCAAATGGTCAGTCGGTTGCGTGTGGTGAATAGACAGTGTATGT 600
Db 541 ACATGTGCTTTCTCTCCAAATGGTCAGTCGGTTGCGTGTGGTGAATAGACAGTGTATGT 600
Qy 601 TCTATCTTTAGCTTAGCTCAACCGCGGACAAAGGATGAACTGTACCGGTTTCAAGAAATG 660
Db 601 TCTATCTTTAGCTTAGCTCAACCGCGGACAAAGGATGAACTGTACCGGTTTCAAGAAATG 660
Qy 661 CTCACGTGTACAGGGATATGTTTCGTGCTGTAGTATGTCCTCCAAATGAGGATGCCAC 720
Db 661 CTCACGTGTACAGGGATATGTTTCGTGCTGTAGTATGTCCTCCAAATGAGGATGCCAC 720
Qy 721 CTTATCACCAAGTTACAGGTGATCAAACTTGATCTTTATGGGATGTAACTACCTGGTCTCAA 780
Db 721 CTTATCACCAAGTTACAGGTGATCAAACTTGATCTTTATGGGATGTAACTACCTGGTCTCAA 780

Qy 781 ACTTCTGTTTTTGGCGGTGAATTTTCAGTCTGGACATATCTGCTGATGATCTAAGCGTCTCA 840
Db 781 ACTTCTGTTTTTGGCGGTGAATTTTCAGTCTGGACATATCTGCTGATGATCTAAGCGTCTCA 840
Qy 841 ATCAGTGGATCAAAACCCAACTGGTTTATATCTGGTTTCATGCGATTTCCACAGCACGGTTG 900
Db 841 ATCAGTGGATCAAAACCCAACTGGTTTATATCTGGTTTCATGCGATTTCCACAGCACGGTTG 900
Qy 901 TGGGACACTCGTGTCTGCAAGCCGAGCAGTGTGCTACCTTTTCATGCTCAGGAGGAGATGTT 960
Db 901 TGGGACACTCGTGTCTGCAAGCCGAGCAGTGTGCTACCTTTTCATGCTCAGGAGGAGATGTT 960
Qy 961 AATACGGTCAAGTTCTTTCCGGATGGGTATAGATTTGGGACTGGATCAGACGATGGAACA 1020
Db 961 AATACGGTCAAGTTCTTTCCGGATGGGTATAGATTTGGGACTGGATCAGACGATGGAACA 1020
Qy 1021 TGCAGGCTGTATGACATAAGGACTGGTCACCACTCCAGGCTCTATCAGGCACATGGTGAT 1080
Db 1021 TGCAGGCTGTATGACATAAGGACTGGTCACCACTCCAGGCTCTATCAGGCACATGGTGAT 1080
Qy 1081 GGTGAGAACCGGACCTGTCACTCCATTTGCATTTCTCTGTGTGAGGAGACTTCTTTTCGCT 1140
Db 1081 GGTGAGAACCGGACCTGTCACTCCATTTGCATTTCTCTGTGTGAGGAGACTTCTTTTCGCT 1140
Qy 1141 GGTATGCGAGCAACAAACATCTTGCTACGTTTGGGATACCTCTTTGGGAGAGGTTGTATTTG 1200
Db 1141 GGTATGCGAGCAACAAACATCTTGCTACGTTTGGGATACCTCTTTGGGAGAGGTTGTATTTG 1200
Qy 1201 GATTTGGGATTTACAGCAGGATTTACACAGGAATAGATTAAGCTGTTTGGGGTGTCTCAGCA 1260
Db 1201 GATTTGGGATTTACAGCAGGATTTACACAGGAATAGATTAAGCTGTTTGGGGTGTCTCAGCA 1260
Qy 1261 GATGGAAGTGCATTTGTGTACAGGAAGTTGGGATTTCAAATCTAAAGATATGGCGCTTTTGA 1320
Db 1261 GATGGAAGTGCATTTGTGTACAGGAAGTTGGGATTTCAAATCTAAAGATATGGCGCTTTTGA 1320
Qy 1321 GGACACAGGAGAGTGTGTAAGAGATTTAAACGAAAAGTAGGAGTCACTCTCCAGTTG 1380
Db 1321 GGACACAGGAGAGTGTGTAAGAGATTTAAACGAAAAGTAGGAGTCACTCTCCAGTTG 1380
Qy 1381 TTGGTTAATATATTTCTGTAGTCGGGAAGTAAAGTTTCGGTTTGTGGAAGGTTTGGTTTGG 1440
Db 1381 TTGGTTAATATATTTCTGTAGTCGGGAAGTAAAGTTTCGGTTTGTGGAAGGTTTGGTTTGG 1440
Qy 1441 AAATAGTGGAGTGGTTAGAAAGAAATTAACCTTCCCTTTTGTGTAGTGTCTTTGATTTATTT 1500
Db 1441 AAATAGTGGAGTGGTTAGAAAGAAATTAACCTTCCCTTTTGTGTAGTGTCTTTGATTTATTT 1500
Qy 1501 ATTTCTTCATTTGGGAATCTAACTCTTCAACACGCTACTCAATGTGAATTTCTGTAAATCAA 1560
Db 1501 ATTTCTTCATTTGGGAATCTAACTCTTCAACACGCTACTCAATGTGAATTTCTGTAAATCAA 1560
Qy 1561 TTGTGTACCCACCACTCTTTACTTACTATCTCTTACTATCTCTTCTATTTGAACGCAAGATAAA 1620
Db 1561 TTGTGTACCCACCACTCTTTACTTACTATCTCTTACTATCTCTTCTATTTGAACGCAAGATAAA 1620
Qy 1621 ACGCTACTA 1629
Db 1621 ACGCTACTA 1629

RESULT 2

AY045860
LOCUS
DEFINITION Arabidopsis thaliana At4g34460 mRNA sequence.
ACCESSION AY045860
VERSION AY045860.1 GI:15028006
KEYWORDS FLU CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsi.

QY	1033	GACATAGGACTGGTGTACCAAACTCCAGGTCTATCAGCCACATGGTGTATGGTGAAGACGGA	1092
Db	1065	GACATAGGACTGGTGTACCAAACTCCAGGTCTATCAGCCACATGGTGTATGGTGAAGACGGA	1124
QY	1093	CCTGTCACTCCTCAATTGCAATCTCTGTGTACAGGAGAGACTTCTTTTTCGCTGGCTATGCGAGC	1152
Db	1125	CCTGTCACTCCTCAATTGCAATCTCTGTGTACAGGAGAGACTTCTTTTTCGCTGGCTATGCGAGC	1184
QY	1153	AACAACTCTTGCTACTAGTTTGGGATACCCCTCTTTGGGAGAGGTTGTATTTGGATTTGGGATTA	1212
Db	1185	AACAACTCTTGCTACTAGTTTGGGATACCCCTCTTTGGGAGAGGTTGTATTTGGATTTGGGATTA	1244
QY	1213	CAGCAGGATTCACACAGGAATAGAACTGTTTGGGGTTGTACAGCAGATGGAAGTGA	1272
Db	1245	CAGCAGGATTCACACAGGAATAGAACTGTTTGGGGTTGTACAGCAGATGGAAGTGC	1304
QY	1273	TTGTGTACAGGAAGTTGGGATTCAAATCTAAAGATATGGCGTTTGGAGGACACAGGAGA	1332
Db	1305	TTGTGTACAGGAAGTTGGGATTCAAATCTAAAGATATGGCGTTTGGAGGAC - CAGGAGA	1363
QY	1333	GTGATTTGAAGAAGATTAAACGAAAGTAGGAGTCAAGTCTCCAGTTGTTGGTTAATATA	1392
Db	1364	GTGATTTGAAGAAGATTAAACG - AAAGTAGGAGTCAAGTCTCCAGTTGTTGGTTAATATA	1422
QY	1393	TTCTGTAGTCGGGAAGTAAGGTTTCGGTTTGTGGAAAGGTTTGGTTTGAATAGTGGAGT	1452
Db	1423	TTCTGTAGTCGGGAAGTAAGGTTTCGGTTTGTGGAAAGGTTTGGTTTGAATAGTGGAGT	1482
QY	1453	GTTTGAAGAATTAAACTTCCCTTTTGTAGTGTCTTTGATTTATTTATTTCTTCATTG	1512
Db	1483	GTTTGAAGAATTAAACTTCCCTTTTGTAGTGTCTTTGATTTATTTATTTCTTCATTG	1542
QY	1513	GGAACTAAACTCTTCAACACGCTACTCAATGTGAATCTGTAACTCAATTTGTACCCAC	1572
Db	1543	GGAACTAAACTCTTCAACACGCTACTCAATGTGAATCTGTAACTCAATTTGTACCCAC	1602
QY	1573	CA 1574	
Db	1603	CA 1604	
RESULT 3			
LOCUS	CQ804428	1134 bp	DNA linear PAT 10-MAY-2004
DEFINITION	Sequence 839 from Patent WO2004035798.		
ACCESSION	CQ804428		
VERSION	CQ804428.1 GI:47110848		
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE	1		
AUTHORS	Inze,D., de Veylder,L. and Vlieghe,K.		
TITLE	Identification of novel e2f target Genes and use thereof		
JOURNAL	Patent: WO 2004035798-A 839 29-APR-2004; Cropdesign N.V. (BE)		
FEATURES	Location/Qualifiers		
source	1..1134		
/organism="Arabidopsis thaliana"			
/mol_type="unassigned DNA"			
/db_xref="taxon:3702"			
ORIGIN			
Query Match	69.5%; Score 1132.4; DB 6; Length 1134;		
Best Local Similarity	99.9%; Pred. No. 2.8e-313;		
Matches 1133; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
QY	208	ATGTCGTCTCCGAGCTCAAGAAGCCACGCCGCTCGCTACCGAGACCGTTAATACCTC	267
Db	1	ATGTCGTCTCCGAGCTCAAGAAGCCACGCCGCTCGCTACCGAGACCGTTAATACCTC	60

QY	268	CGTGAACAGCTTTAGACAGAGACGCTCCAGTCTCTCGATACCGATGTGGCGAGGTATTCA	327
Db	61	CGTGAACAGCTTTAGACAGAGACGCTCCAGTCTCTCGATACCGATGTGGCGAGGTATTCA	120
QY	328	CGCGCGAAGACCTACTCGGGTGAGCTTCGGAGCAACGGATCTGGTTGTTCGTA	387
Db	121	CGCGCGAAGACCTACTCGGGTGAGCTTCGGAGCAACGGATCTGGTTGTTCGTA	180
QY	388	CTTCAGGACACACCGGAAAGGTTTATTCATTAGATTGGACACCGGAGAGGAAACCGGATT	447
Db	181	CTTCAGGACACACCGGAAAGGTTTATTCATTAGATTGGACACCGGAGAGGAAACCGGATT	240
QY	448	GTCAAGTGCATCTCAAGATGGGAGATTAAATCGTGTGGAAATGCTCTAAACGAGTCAGAAACT	507
Db	241	GTCAAGTGCATCTCAAGATGGGAGATTAAATCGTGTGGAAATGCTCTAAACGAGTCAGAAACT	300
QY	508	CATGCTATTAACCTCCCTTGTGTCATGGGTTATGACATGTGCTTCTCTCCAAATGGTCAG	567
Db	301	CATGCTATTAACCTCCCTTGTGTCATGGGTTATGACATGTGCTTCTCTCCAAATGGTCAG	360
QY	568	TCGGTTGCGTGTGGATTAGACAGATGTATGTTCTATCTTTAGCCTTAGCTCAACCGCG	627
Db	361	TCGGTTGCGTGTGGATTAGACAGATGTATGTTCTATCTTTAGCCTTAGCTCAACCGCG	420
QY	628	GACAAAGATGGAATCTGACCGGTTTCAAGAAATGCTCTCTGCTCACAGGGAGATATGTTTCG	687
Db	421	GACAAAGATGGAATCTGACCGGTTTCAAGAAATGCTCTCTGCTCACAGGGAGATATGTTTCG	480
QY	688	TGCTGTCAAGTATGCCAAATGAGGATGCCACCTTATCACAGTTCAGGTCAGTCAAACT	747
Db	481	TGCTGTCAAGTATGCCAAATGAGGATGCCACCTTATCACAGTTCAGGTCAGTCAAACT	540
QY	748	TGTATCTTATGGGATGTAACTACTGGTCTCAAAACTCTCTGTTTGGCGGTGAATTTTCAG	807
Db	541	TGTATCTTATGGGATGTAACTACTGGTCTCAAAACTCTCTGTTTGGCGGTGAATTTTCAG	600
QY	808	TCTGGAACATCTGCTGATGTACTAAGCGTCTCAATCATGATGGATCAAAACCCAACTGGTTT	867
Db	601	TCTGGAACATCTGCTGATGTACTAAGCGTCTCAATCATGATGGATCAAAACCCAACTGGTTT	660
QY	868	ATATCTGTTTCATCGGATTTCCACAGCAGCGTGTGGGACACTCGTGTGCAAGCCGACGA	927
Db	661	ATATCTGTTTCATCGGATTTCCACAGCAGCGTGTGGGACACTCGTGTGCAAGCCGACGA	720
QY	928	GTGCGTACCTTTTCATGGTTCAGGAGGAGATGTTAATACGGTCAAGTCTCTTTCCGATGGG	987
Db	721	GTGCGTACCTTTTCATGGTTCAGGAGGAGATGTTAATACGGTCAAGTCTCTTTCCGATGGG	780
QY	988	TATAGATTGGGACTGGATCAGAGATGGAAACATGCGGCTGTATGACATAGGACTGGT	1047
Db	781	TATAGATTGGGACTGGATCAGAGATGGAAACATGCGGCTGTATGACATAGGACTGGT	840
QY	1048	CACCACTCCAGGTCATACGCCACATGGTGTGAGAACGGACCTGTCCACCTCCATT	1107
Db	841	CACCACTCCAGGTCATACGCCACATGGTGTGAGAACGGACCTGTCCACCTCCATT	900
QY	1108	GCATTTCTGTGTACGGGAGACTTTCTTTTCGCTGGCTATGCGAGCAACAACTTGTCTAC	1167
Db	901	GCATTTCTGTGTACGGGAGACTTTCTTTTCGCTGGCTATGCGAGCAACAACTTGTCTAC	960
QY	1168	GTTTGGGATACCTCTTGGGAGAGGTTGTATGGATTGGGATTTACAGCAGGATTCACAC	1227
Db	961	GTTTGGGATACCTCTTGGGAGAGGTTGTATGGATTGGGATTTACAGCAGGATTCACAC	1020
QY	1228	AGGAATAGAAATAGCTGTTTGGGCTGTGCACAGATGGAACTGTCATGTCACAGGAGT	1287
Db	1021	AGGAATAGAAATAGCTGTTTGGGCTGTGCACAGATGGAACTGTCATGTCACAGGAGT	1080
QY	1288	TGGGATTCAAATCTAAAGATATGGCGTTTGGAGGACACAGGAGAGTGAATTGA	1341
Db	1081	TGGGATTCAAATCTAAAGATATGGCGTTTGGAGGACACAGGAGAGTGAATTGA	1134

```

RESULT 4
BT020366
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

    1128 bp mRNA linear PLN 04-DEC-2004
    Arabidopsis thaliana At4g34460 gene, complete cds.
    BT020366.1 GI:56382004
    FLI CDNA.
    Arabidopsis thaliana (thale cress)
    Arabidopsis thaliana
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
    rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
    1 (bases 1 to 1128)
    Kim, C.J., Chen, H., Cheuk, R., Shinn, P. and Ecker, J.R.
    Arabidopsis ORF clones
    Unpublished
    2 (bases 1 to 1128)
    Kim, C.J., Chen, H., Cheuk, R., Shinn, P. and Ecker, J.R.
    Direct Submission
    Submitted (04-DEC-2004) Salk Institute Genomic Analysis Laboratory
    (SIGNAL), Plant Biology Laboratory, The Salk Institute for
    Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
    USA
    The discrepancy does not affect the protein sequence.
    Possible splice variant.
    Location/Qualifiers
    1. .1128
    /organism="Arabidopsis thaliana"
    /mol_type="mRNA"
    /db_xref="taxon:3702"
    /chromosome="4"
    /clone="U68637"
    /ecotype="Columbia"
    /note="This clone is in pUNI 51"
    1. .1128
    /note="unknown protein"
    /codon_start=1
    /product="At4g34460"
    /protein_id="AAV85721.1"
    /db_xref="GI:56382005"
    /translation="MSVSLKERHVAETVNNLRDLPORLQLLDTVARVSAQG
    RRVSGADLVCCRLQHTGKYSVLDTWPERNIVSASQGRLLVNNALTSQKTHA
    IKLPANWMTAFSPNGSVACGLDSVCSIFSLSTADKDGTPVSRMLTGHGTVS
    CQYVNEADHLITSSGDDTCILMDVTGLKTSVFGEFQSGHTADVLVSISGNPN
    WFTSGSDTARLMDTRAAARVTPFHGEGDVNTVKFPDGRFQSGDDGTCRLYD
    IRTGHOLQVYOPHGDGNGPVTSLAFSVSGRLLFAGYASNNTCYVMDTLGEVLDLG
    LQDSHRNRISCLGLSADGSALCTGSWDSNLKIWAFGGPGE"
    misc_difference 927
    /note="compared to genome"
    /replace="t"

ORIGIN
Query Match 68.4%; Score 1113.8; DB 15; Length 1128;
Best Local Similarity 99.8%; Pred. No. 8.66-308;
Matches 1115; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 208 ATGCTGTCTCCGAGCTCAAGAACCCGACCGCTCGCTACGAGACCGGTAAATAACCTC 267
DB 1 ATGCTGTCTCCGAGCTCAAGAACCCGACCGCTCGCTACGAGACCGGTAAATAACCTC 60

QY 268 CGTGACAGCTTAGACAGACAGCCCTCCAGCTCCTCGATACCGATGTGGCGAGGTATCA 327
DB 61 CGTGACAGCTTAGACAGACAGCCCTCCAGCTCCTCGATACCGATGTGGCGAGGTATCA 120

QY 328 GCGGCGCAGGACGTACTCGGCTCGGAGTTCGGAGCAACCGATCTGGTTGTCCTACT 387
DB 121 GCGGCGCAGGACGTACTCGGCTCGGAGTTCGGAGCAACCGATCTGGTTGTCCTACT 180

QY 388 CTTTCAGGACACACCGGAAAGGTTTATTTCATTAGATTGGACACCGGAGGAAACCGGATT 447
DB 181 CTTTCAGGACACACCGGAAAGGTTTATTTCATTAGATTGGACACCGGAGGAAACCGGATT 240

```

```

QY 448 GTCAGTGCATCTCAAGATGGGAGATTAAATCGTGTGGAAATGCTTAACGAGTCAGAAAAC 507
DB 241 GTCAGTGCATCTCAAGATGGGAGATTAAATCGTGTGGAAATGCTTAACGAGTCAGAAAAC 300
QY 508 CATGCTATTAAACTCCCTTGTGCAATGGTATTATGATGCTTCTCTCCAAATGGTCAG 567
DB 301 CATGCTATTAAACTCCCTTGTGCAATGGTATTATGATGCTTCTCTCCAAATGGTCAG 360
QY 568 TCGGTTGGCTGTGGATTAGACAGATGATGTTCTATCTTTAGCTTTAGCTCAACGGCG 627
DB 361 TCGGTTGGCTGTGGATTAGACAGATGATGTTCTATCTTTAGCTTTAGCTCAACGGCG 420
QY 628 GACAGGATGCAACTGTACCGGTTTCAAGATGCTCACTGTCACAGGGGATATGTTTCG 687
DB 421 GACAGGATGCAACTGTACCGGTTTCAAGATGCTCACTGTCACAGGGGATATGTTTCG 480
QY 688 TGCTGTCTAGTATGTCCTCAAAATGAGGATGCCACCTTATCACAGTTCAGGTGATCAAACT 747
DB 481 TGCTGTCTAGTATGTCCTCAAAATGAGGATGCCACCTTATCACAGTTCAGGTGATCAAACT 540
QY 748 TGTATCTTATGGGATGTAACCTGCTCAAACTTCTGTTTGGCGGTGAATTTTCAG 807
DB 541 TGTATCTTATGGGATGTAACCTGCTCAAACTTCTGTTTGGCGGTGAATTTTCAG 600
QY 808 TCTGACATACCTGCTGATGATGTAAGCTCTCAATCAGTGTGATCAAACTCAAACTGGTTT 867
DB 601 TCTGACATACCTGCTGATGATGTAAGCTCTCAATCAGTGTGATCAAACTCAAACTGGTTT 560
QY 868 ATATCTGCTTATGCGATTCCACAGCACCGGTTGTGGGACACTCGTGTGTCGAAGCGGCA 927
DB 661 ATATCTGCTTATGCGATTCCACAGCACCGGTTGTGGGACACTCGTGTGTCGAAGCGGCA 720
QY 928 GTGCGTACCTTTCTGTCACGAGGAGATGTTTAATACGTCAGTTCCTTCCGATGGG 987
DB 721 GTGCGTACCTTTCTGTCACGAGGAGATGTTTAATACGTCAGTTCCTTCCGATGGG 780
QY 988 TATAGATTGGGACCTGGATCAGACGATGEMACATGACGCTGTATGACATAAGGACTGTT 1047
DB 781 TATAGATTGGGACCTGGATCAGACGATGEMACATGACGCTGTATGACATAAGGACTGTT 840
QY 1048 CACCAACTCCAGGCTCTATCAGCCACATGCTGTGATGTTGAGAACGGACCTGTCTACCTCCATT 1107
DB 841 CACCAACTCCAGGCTCTATCAGCCACATGCTGTGATGTTGAGAACGGACCTGTCTACCTCCATT 900
QY 1108 GCAATCTCTGTCAGGAGACTTCTTTTGGCTGCTATCGAGCAACAACTTGTCTAC 1167
DB 901 GCAATCTCTGTCAGGAGACTTCTTTTGGCTGCTATCGAGCAACAACTTGTCTAC 960
QY 1168 GTTTCGGATACCTCTTGGGAGAGCTTGTATTGGATTGGGATTACAGCAGGATTCACAC 1227
DB 961 GTTTCGGATACCTCTTGGGAGAGCTTGTATTGGATTGGGATTACAGCAGGATTCACAC 1020
QY 1228 AGGAATAGAAATAAGCTGTTTGGGGTTGTGTCAGCAGATGGAAGTGCATTGTTGTACAGAACT 1287
DB 1021 AGGAATAGAAATAAGCTGTTTGGGGTTGTGTCAGCAGATGGAAGTGCATTGTTGTACAGAACT 1080
QY 1288 TGGGATTCAAAATCTAAAGATATGGCGTTTGGAGGAC 1324
DB 1081 TGGGATTCAAAATCTAAAGATATGGCGTTTGGAGGAC 1117

```

```

RESULT 5
ATT4L20/c
LOCUS
DEFINITION
    Arabidopsis thaliana DNA chromosome 4, BAC clone T4L20 (ESSA
    project).
    AL023094
    AL023094.2 GI:5678625
    Arabidopsis thaliana (thale cress)
    Arabidopsis thaliana
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

```


11372. .11569, .11651. .11795, .11892. .12046, .12237. .12476,
12610. .12804)
/gene="T4L20.30"
/note="strong similarity to coat protein gamma-cop, Bos
primigenius
Contains 2-oxo acid dehydrogenases acyltransferase
component lipoyl binding site, Lipoyl
[NPVSSAALVSGHLKTNQIVRWSNEV]
contains EST gd:AA395649, N96014, H36940, R90482,
AA042453, T43773, R90315, T46306, T75984"
/codon_start=1
/product="Noncathrin coat protein gamma-like protein"
/protein_id="CAA18824.1"
/db_xref="GI:3641837"
/db_xref="GOA:O65673"
/db_xref="InterPro:IPR002553"
/db_xref="InterPro:IPR03016"
/translation="MAQPLKKDDHDDDELEYSPFMGIEKGAQLVEARVFNPDQVDR
RCSQVITKLILNNGESFTEATEVFSVTKLQSQDGTGLRMVYLIIEKSPSSD
EVIYTSISLMKMSKIDMYRANAI RVLRCITDGLLTQIERYLKQAI VDKNPVSSA
ALVSGHLKTNQIVRWSNEVEGIQSRSAVQFIALHQRNDRLAVSKLGV
STRCGVSPLQAQCLIQRPVEFLESLCHKAEMVILAEARATFELDGVTSRELTPA
ITVLQPISSFPVLRFPAVATLKNVAMTHPMATVNCIDMESLISQNRSLATLAI
TLTKGTGNESSVERLMQVITNFMSSDADEFKIVVDAIRSLCVKFKPLKTRSLMTPLSNI
LKEEGFGRKAIADVISITITRDPAKESGLLHCFIEDCEFTYILSTQILHFLGIE
DNDSDPKYIRIYINRVHLNATVRAAAVSTLAKGFPMVBSLKPRTIVLLKRCIYDS
GENTEDRATLYLISEPSEAFDINSVPKEVKSQPLAEKQAQKPTGLGAPPAAPASGF
DGYERLLGSIPEFAAFGKLKSSLPVELTEAEVAVNVVGHIFDSHVVFQYCNNTITV
PEOLLERVLNIEAEFSFSTKALNSLPYDSGQAAVVFPEKPGAVAGKFSNTLTIV
VKEVHVDSTGEADDGVEDEYQLEDLEWAGDYMVKGVSNFRNANESMDDEDERVD
EYGLGDSGLGAVKAVMDLGMQTCGETTIFLNARSHCTLLSGVTIGNVKVLVRAQ
FGMGSQDIAMKLVRAEDSVSAEIHIVAS"
7965. .8010
/gene="T4L20.30"
/number=1
8011. .8234
/gene="T4L20.30"
/number=1
8235. .8383
/gene="T4L20.30"
/number=2
8384. .8485
/gene="T4L20.30"

Query Match	49.8%;	Score 811.8;	DB 15;	Length 125502;
Best Local Similarity	68.5%;	Pred. No. 5.3e-221;		
Matches 1614; Conservative	0;	Mismatches 2;	Indels 741;	Gaps 6;

Qy	13	CGTGGTTGTGCTTGACTGATCTCTCTCAAGCTTTTTTAATCTCTCTCTCTTTTCCTCA 72
Db	15470	CGTGGTTGTGCTTGACTGATCTCTCTCAAGCTTTTTTAATCTCTCTCTCTTTTCCTCA 15411
Qy	73	CGTAAATCCGCCAAAATCCCATCTTTCTAGGGTTCGATCTCCCCTCTCTCAAATCATGAACCT 132
Db	15410	CGTAAATCCGCCAAAATCCCATCTTTCTAGGGTTCGATCTCCCCTCTCTCAAATCATGAACCT 15351
Qy	133	TCTTCTCTCTAGACCCCAACAAGTTTCCCCCTTTATTGTATCGGCACGGAGAAGCT 192
Db	15350	TCTTCTCTCTAGACCCCAACAAGTTTCCCCCTTTATTGTATCGGCACGGAGAAGCT 15291
Qy	193	AAGCTCTGATCCCGGAATGCTGTCTCCGAGCTCAAAGAACGCCCGCTCGCTACGGAG 252
Db	15290	AAGCTCTGATCCCGGAATGCTGTCTCCGAGCTCAAAGAACGCCCGCTCGCTACGGAG 15231
Qy	253	ACGTTTAATAAACCCTCCGTGACACGATTAGACAGAGACGCTCCAGCTCCTCGATACCGAT 312
Db	15230	ACCGTTTAATAAACCCTCCGTGACACGATTAGACAGAGACGCTCCAGCTCCTCGATACCGAT 15171
Qy	313	----- 312
Db	15170	GGTTGTGGATTCCTTTTCGTCTCTCTCTCTCTCATCTCTCTCTCTTTTAATTGGAT 15111

313	----	GTGGCAGGTATTCTACGGCGCAGGACGTACTCGGGTGAGCTTCGGA	360
15110	TGAATTTTTCAGTGGCGAGGTATTTACGGCGCGCAAGGACGTACTCGGGTGAGCTTCGGA	15051	
361	GCACCGGATCTGGTTTGTGTGCTACTCTTCAGGGACACACCGGAA	406	
15050	GCACCGGATCTGGTTTGTGTGCTACTCTTCAGGGACACACCGGAAAGGTGACTCAAAGT	14991	
407	-----	406	
14990	TTGCATCTTTATCCTTCAAGCATGCTCTTTGTGGAGTGTTTCTTAGCCTGTGATAAAA	14931	
407	-----	406	
14930	GCTTCCATTTTGGGGTGTATCTAGTTTGTATTTTAGTTGGTGTGTTTATAAGATGTTA	14871	
407	-----	406	
14870	CCAGTTATCATGTATGAGGATTTCGTTAAGNCTCTACGGAAATCTAGACAGAGAGTT	14811	
407	-----	421	
14810	GATTGTGAATGTTAGCTTGTGTTGATTTAACTATGTGGGATTTGCAGGTTATTTCAATTAG	14751	
422	ATTGACACACCGAGAGGAAACCGGATTTGCAGTGATCTCAAGATGGGAGATTTAATCGTGT	481	
14750	ATTGACACACCGAGAGGAAACCGGATTTGCAGTGATCTCAAGATGGGAGATTTAATCGTGT	14691	
482	GGAAATGCTCTAACAGTACAGAAACTCATGCTATTAACCTCCCTTTGTCATGGGTTATGA	541	
14690	GGAAATGCTCTAACAGTACAGAAACTCATGCTATTTAACTCCCTTTGTCATGGGTTATGA	14631	
542	CATGTGCTTTCTCTCCAAATGGTCAGTCCGTTGGTGTGGATTAGACAGTGATATGTT	601	
14630	CATGTGCTTTCTCTCCAAATGGTCAGTCCGTTGGTGTGGATTAGACAGTGATATGTT	14571	
602	CTATCTTTAGCCTTAGCTCAACCGCGGACAGGATGGAACTGTATCCGGTTTCAAGAATGC	661	
14570	CTATCTTTAGCCTTAGCTCAACCGCGGACAGGATGGAACTGTATCCGGTTTCAAGAATGC	14511	
662	TCATGCTCACAGGGGATATGTTTCGTGCTCTCAGTATGTCCTCAAAATGAGGATGCCAC	721	
14510	TCATGCTCACAGGGGATATGTTTCGTGCTCTCAGTATGTCCTCAAAATGAGGATGCCAC	14451	
722	TTATCACCAAGTTTCAGGTGATCAAACTTGATCTTATGGGATGTAACTACTGGTCTCAAAA	781	
14450	TTATCACCAAGTTTCAGGTGATCAAACTTGATCTTATGGGATGTAACTACTGGTCTCAAAA	14391	
782	CTTCTGTTTTTGGCGGTGAATTCAGTCTGACATATCTGCTGATGTACTA	831	
14390	CTTCTGTTTTTGGCGGTGAATTCAGTCTGACATATCTGCTGATGTACTAAGGTATGGAT	14331	
832	-----	831	
14330	ATTGCTTTTCATCTCCTTGGCTCATTTGGCATCAGTTTGTGATTAAGGTAGAATATATTTG	14271	
832	-----	831	
14270	AAAACTACTTAATCTTCCAGAACCATACCATTGGATCGAGTGTGGAAATCAACATATA	14211	
832	-----	831	
14210	GGTTTGGTCTTATTGAATAGATACGTTTATTGACTTGTGTATAAAAGCTAGCATGATG	14151	
832	-----	AG 833	
14150	ATATATTGCAAGTCAGTTAAACTCCGCTCTTACTTGTGTGATTAATTCGCCAATCTTGAG	14091	
834	CGTCTCAATCAGTGGATCAAAACCAAACTGGTTTATATCTGTTTCATCGGATTCACAGC	893	
14090	CGTCTCAATCAGTGGATCAAAACCAAACTGGTTTATATCTGTTTCATCGGATTCACAGC	14031	
894	ACGGTTGTGGGACACTCGTGTCTGACGCGGACGAGTGGGTACCTTTCTATGGTACGAGGG	953	

```
|||||
14030 ACGGTGTGGGACACTGCTGCAAGCGACAGTGGTACCTTTCATGGTCAGAGG 13971
QY 954 AGATGTTAATACGGTCAAGTCTCTTCCGATGGGTATAGATTGGGACTGGATCAGACGA 1013
Db AGATGTTAATACGGTCAAGTCTCTTCCGATGGGTATAGATTGGGACTGGATCAGACGA 13911
QY 1014 TGGACATGCAGGCTGTATGACATGAAGGACTGGTCAACCACTCCAGGCTCATCAGCCACA 1073
Db TGGACATGCAGGCTGTATGACATGAAGGACTGGTCAACCACTCCAGGCTCATCAGCCACA 13851
QY 1074 TGGTCAATGTCGAGACGACCTGTCACCTCCATTCCTGCTGTCAGGAGACTTCT 1133
Db TGGTCAATGTCGAGACGACCTGTCACCTCCATTCCTGCTGTCAGGAGACTTCT 13791
QY 1134 TTTCCGCTGGCTATGCGAGCAACAACTGCTGTCACCTTTGGGATACCTCTTGGGAG--- 1189
Db TTTCCGCTGGCTATGCGAGCAACAACTGCTGTCACCTTTGGGATACCTCTTGGGAGAGGT 13731
QY 1190 ----- 1189
Db 13730 AAGCAATTAGTATTGAGTAATGAGTGGTGGAGTGTGTTTCGCTCATGTAAGGTTTATG 13671
QY 1190 -----AGTTGTATTGGATTTGGGATTTAGGATTTGGGATTTAGCAGGATTCA 1224
Db 13670 ATTGAAACGTGAGGAGTTCCTTCGACAGTGTGATTGGATTTGGGATTTAGCAGGATTCA 13611
QY 1225 CACAGGAATAGAAATAGCTGTTGGGGTGTGTCAGCAGATGGAAGTGCATTGTGTACAGGA 1284
Db 13610 CACAGGAATAGAAATAGCTGTTGGGGTGTGTCAGCAGATGGAAGTGCATTGTGTACAGGA 13551
QY 1285 AGTTGGGATTCAAATCTAA----- 1303
Db 13550 AGTTGGGATTCAAATCTAAAGGTTAGTACTTAGTAGTGTCTCTGCTTATCATTTACCATG 13491
QY 1304 -----AGATATCGG 1312
Db 13490 TTATGTTTGAAGGAAATTAGAGGATAGATGAATTCGTTGAATTCACAGATATCGG 13431
QY 1313 CGTTTGGAGGACACAGGAGAGTGATTGAAGAAGATTTAAACGAAAGTAGGAGTCAACGTC 1372
Db 13430 CGTTTGGAGGACACAGGAGAGTGATTGAAGAAGATTTAAACG- AAGTAGGAGTCAACGTC 13372
QY 1373 TCCAGTTGTTGGTAAATATATCTGTAGTCGGGAAGTAAAGGTTTCGGTTTGTGGAAGGTGT 1432
Db 13371 TCCAGTTGTTGGTAAATATATCTGTAGTCGGGAAGTAAAGGTTTCGGTTTGTGGAAGGTGT 13312
QY 1433 TTGGTTTCAAAATAGTGGAGTGTGTAAGAAGATTAACCTTCCCTTTTCTGTAGTGTGCTTTG 1492
Db 13311 TTGGTTTCAAAATAGTGGAGTGTGTAAGAAGATTAACCTTCCCTTTTCTGTAGTGTGCTTTG 13252
QY 1493 ATTTATTTATTTCTTCAATGGGAATAAAGTCTTCAACAGGCTACTCAATGTGAATTCCT 1552
Db 13251 ATTTATTTATTTCTTCAATGGGAATAAAGTCTTCAACAGGCTACTCAATGTGAATTCCT 13192
QY 1553 GTAATCAATGTGTACCAACAGGCTTTTACTTATCATCTCTTCAATGTGAACGAG 1612
Db 13191 GTAATCAATGTGTACCAACAGGCTTTTACTTATCATCTCTTCAATGTGAACGAG 13132
QY 1613 AAGATAAACCCTACTA 1629
Db 13131 ACGATAAACCCTACTA 13115
```

RESULT 6

```
LOCUS ATCHRIV81/c 195921 bp DNA linear PLN 16-APR-2005
DEFINITION Arabidopsis thaliana DNA chromosome 4, contig fragment No. 81.
ACCESSION AL161585
VERSION AL161585.2 GI:7270366
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 36279)

Weichselgartner,M., Fartmann,B., Granderath,K., Dauner,D., Herzl,A., Neumann,S., Mewes,H.W., Lemcke,K. and Mayer,K.F.X. Unpublished

2 (bases 8520 to 8748)

Volckaert,G., Grymonprez,B., Voet,M., Robben,J., Mewes,H.W., Lemcke,K. and Mayer,K.F.X. Unpublished

3 (bases 109690 to 195921)

Terry,N., Ardiles,W., Buysschaert,C., Dasseville,R., De Clerck,R., De Keyser,A., Rouze,P., Veyt,P., Van Den Daele,H., Villaroel,R., Gielen,J., Van Montagu,M., Mewes,H.W., Lemcke,K. and Mayer,K.F.X. Unpublished

4 (bases 1 to 195921)

EU Arabidopsis sequencing,project. Direct Submission

Submitted (10-MAR-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de,project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk

Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/> this fragment has an overlap with ATCHRIV80 at the 5' end and an overlap with ATCHRIV82 at the 3' end.

FEATURES

Location/Qualifiers

1..195921

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/chromosome="4"

/ecotype="Columbia"

8520..36287

/note="position 1-27768 overlaps to BAC F28A23, EMBL acc AL021961; please refer to this entry for analysis"

13783..14721

/gene="AT4g34170"

join(13783..13980,14038..14721)

/gene="AT4g34170"

/note="similarity to Human mRNA for KIAA0132 gene, PID:D1010123

contains EST gb:AI992809.1, Z34043, H77079, T20584"

/codon_start=1

/product="putative protein"

/protein_id="CAB80134.1"

/db_xref="GI:7270367"

/db_xref="InterPro:IPR001810"

/db_xref="InterPro:IPR006652"

/db_xref="InterPro:IPR011498"

/db_xref="UniProt/TrEMBL:O49488"

/translation="MDDGEELPVKTMMLHDDLIINCLARVSRSNHPTLSLVCKRPHS LLASVELYQRTLLTGERCFYQYSSRKILVQLSPNSTSAGTAVVGNDAICGGI KNTLSSVMYDNRSHTRBAPSRVPRMPSPVCTLDGKIYVMGCDNLDSNWEVF DTKQTWEFLQIPSEELPGGSAYVECTVYVMSSEKDVTVKLHGRWSAADSAN GWGPGSSYCVIENLVSLCFVHKIRWYDKRWVTPKLGFLPSLPCNGHVKLDYGEKM VILWEKYVDVDEKKMIW"

13783..13980

/gene="AT4g34170"

/number=1

13981..14037

/gene="AT4g34170"

/number=1

14038..14721

/gene="AT4g34170"

/number=2

14960..17079

/gene="AT4g34180"

complement(join(14960..15200,15752..15972,16074..16197,


```

16443. .16515,16601. .16645,16725. .17079))
/ gene="AT4g34180"
complement(join(14960. .15200,15752. .15972,16074. .16197,
16443. .16515,16601. .16645,16725. .17079))
/ gene="AT4g34180"
/ note="Similarity to hypothetical protein slr2121,
Synecocystis sp. PIR2:875497
contains EST gb:N64521, T45635, T45634, T21897, N38394,
T44783, T46555, H37721, A1993997.1"
/ codon_start=1
/ product="putative protein"
/ protein_id="CAB80135.1"
/ db_xref="GI:7270368"
/ db_xref="InterPro:IPR007325"
/ db_xref="UniProt/TREMBL:O49487"
/ translation="MTRSVSPFLFPAVLISLSSLLADDPKPIRREYVGGKIYDIS
HRYTPEIPAWESSGLGKTLRLAASMKNGSFANVSEKLSVHSGTHVDPAGHFWQNY
YDAGFDSDLDLQVLNGPALLVDVPRDKNTAEVMSLIHQGVRRVLFRTSTNDLY
YTFSEGYSLNPNGLDYLSPAAPESPATHRVILKGRAYSLLDKKREIGRLKRENN
YVLTREIHPFSGYNSAAGWCGGRNLLASLTAISWSGRSTDQKHSHOVIOGKR
KHNLCVLVWFHYIERDQVPSSLQKASVLTKEKPLNAQESVLVHQLNDNRSEPAL
TTKRLLDVAKRIAKERS"
complement(14960. .15200)
/ gene="AT4g34180"
/ number=1
complement(15201. .15751)
/ gene="AT4g34180"
/ number=1
complement(15752. .15972)
/ gene="AT4g34180"
/ number=2
complement(15973. .16073)
/ gene="AT4g34180"
/ number=2
complement(16074. .16197)
/ gene="AT4g34180"
/ number=3
complement(16198. .16442)
/ gene="AT4g34180"
/ number=3
complement(16443. .16515)
/ gene="AT4g34180"
/ number=4
complement(16516. .16600)
/ gene="AT4g34180"
/ number=4
complement(16601. .16645)
/ gene="AT4g34180"
/ number=5
complement(16646. .16724)
/ gene="AT4g34180"
/ number=5
complement(16725. .17079)
/ gene="AT4g34180"
/ number=6
18302. .19070
/ gene="AT4g34190"
complement(join(18302. .18406,18513. .18691,18950. .19070))
/ gene="AT4g34190"
complement(join(18302. .18406,18513. .18691,18950. .19070))
/ gene="AT4g34190"
/ note="contains EST gb:A1995995.1, A3395832, A3394538,
225494, T45229, N65188, Z26670, R86822"
/ codon_start=1
/ product="putative protein"
/ protein_id="CAB80136.1"
/ db_xref="GI:7270369"
/ db_xref="UniProt/TREMBL:O49486"
/ translation="MFPDSGALKLATITNPSTCRVHPQLAGIRSTFASGSLPLK
LSMTRGGNRAASVIRSEQSGSLDIWLGRGMVGFVAITVEISTGKGLLENF
GVASPLPTVALVGLVAAVFIQSSKN"
complement(18302. .18406)
/ gene="AT4g34190"

```

```

/number=1
complement(18407. .18512)
/ gene="AT4g34190"
/number=1
complement(18513. .18691)
/ gene="AT4g34190"
/number=2
complement(18692. .18949)
/ gene="AT4g34190"
/number=2
complement(18950. .19070)
/ gene="AT4g34190"
/number=3
19737. .22257
/ gene="AT4g34200"
complement(join(19737. .20408,20757. .20852,20943. .21430,
21702. .22257))
/ gene="AT4g34200"
complement(join(19737. .20408,20757. .20852,20943. .21430,
21702. .22257))
/ gene="AT4g34200"
/ notes="strong similarity to Phosphoglycerate
dehydrogenase, Arabidopsis thaliana, PATCHX:D1021238
Contains ATP/GTP-binding site motif A (P-loop)
AA195-202;D-isomer specific 2-hydroxyacid dehydrogenases
signatures AA203-230;D-isomer specific 2-hydroxyacid
dehydrogenases signatures AA251-273;D-isomer specific
2-hydroxyacid dehydrogenases signatures AA280-296
contains EST gb:A1996638.1, AA585787, Z34025, Z34777,
T22482, W43188"
/ codon_start=1
/ product="phosphoglycerate dehydrogenase-like protein"
/ protein_id="CAB80137.1"
/ db_xref="GI:7270370"
/ db_xref="GOA:O49485"
/ db_xref="InterPro:IPR00205"
/ db_xref="InterPro:IPR002912"
/ db_xref="InterPro:IPR006139"
/ db_xref="InterPro:IPR006140"
/ db_xref="InterPro:IPR006236"
/ db_xref="UniProt/TREMBL:O49485"
/ translation="MSATAAASSIAVATNSLRNVLTSSRPLSPALSIVAPSSGRNT
LQRLVLVSGSTGDSKPTILVAEKLGDAGIKLEEDVANDVCSYNTPEELNIKISLC
DALIVRSCTKVGREVPFESSGRULKVCGRAGVIGINDVLSAETGECVLVNPATNTIA
AAEGIALMAAMANNVAQASDASVAGAKRKNKVGSLVKTLLAVLFGKVGTEVARR
AKGLMRVIAHDYPADRAHAIGVDLVDFEALATADFISLHMLPTTTSKILNDET
FAKMKGVRIWNVARGGVIDEDALVALDAGIVAQAALDVFTKEPPAKDSKLQHERV"
Query Match 49.8%; Score 811.8; DB 15; Length 195921;
Best Local Similarity 68.5%; Pred. No. 5.5e-221;
Matches 1614; Conservative 0; Mismatches 2; Indels 741; Gaps 6;
QY 13 CGTGTGTTGTCTTGACTGATTTCTCTCAAGCTTTTTTAATCTCTCTCTTTTCCCA 72
DB 125159 CGTGTGTTGTCTTGACTGATTTCTCTCAAGCTTTTTTAATCTCTCTCTTTTCCCA 125100
QY 73 CGTAATTCCTCCCAATTCATTTCTTAGGTTTCGATTCCTCTCTCAATCATGAACCT 132
DB 125099 CGTAATTCCTCCCAATTCATTTCTTAGGTTTCGATTCCTCTCTCAATCATGAACCT 125040
QY 133 TCTTCTCTTCTAGACCCACAAAGTTTCCCTTTTATTTTCATCGGCGACGAGAAGCT 192
DB 125039 TCTTCTCTTCTAGACCCACAAAGTTTCCCTTTTATTTTCATCGGCGACGAGAAGCT 124980
QY 193 AAGTCTGATCCCGGAATGCTGCTCTCCGAGCTCAAAGAACCCACGCCGCTCGTACGGAG 252
DB 124979 AAGTCTGATCCCGGAATGCTGCTCTCCGAGCTCAAAGAACCCACGCCGCTCGTACGGAG 124920
QY 253 ACCGTTAATAACCTCCGTTAGACAGCTTAGACAGAGACGCTTCCAGCTCTCGATACCGAT 312
DB 124919 ACCGTTAATAACCTCCGTTAGACAGCTTAGACAGAGACGCTTCCAGCTCTCGATACCGAT 124860
QY 313 ----- 312

```

[illegible]

Db	123779	CGTCTCAATCAGTGGATCAAAACCCAACTGGTTTTATATCTCGTTTCATCGGATTCACAGC	123782
Qy	894	ACGGTTGTGGGACACTCGTGTCTCAAGCCGAGCAGTGCGTACCTTTTCATGGTCACGAGGG	953
Db	123719	ACGGTTGTGGGACACTCGTGTCTCAAGCCGAGCAGTGCGTACCTTTTCATGGTCACGAGGG	123660
Qy	954	AGATGTTTAATACGGTCAAGTTCITTTCCGGATGGGTATAGATTTGGGACTGGATCAGACGA	1013
Db	123659	AGATGTTTAATACGGTCAAGTTCITTTCCGGATGGGTATAGATTTGGGACTGGATCAGACGA	123600
Qy	1014	TGGAACATGCAAGCTGTATGACATAAAGGACTGGTCACCAACTCCAGGTCCTATCAGCCACA	1073
Db	123599	TGGAACATGCAAGCTGTATGACATAAAGGACTGGTCACCAACTCCAGGTCCTATCAGCCACA	123540
Qy	1074	TGGTGATGGTGAGAACCGACCTGTCACTCCATTTGCAATTCCTGTGTGTCAGGAGACTTCT	1133
Db	123539	TGGTGATGGTGAGAACCGACCTGTCACTCCATTTGCAATTCCTGTGTGTCAGGAGACTTCT	123480
Qy	1134	TTTTCGCTGGCTATCGGAGCAACAACACTTGTCTACGTTTGGGATACCTCTTTGGGAG---	1189
Db	123479	TTTTCGCTGGCTATCGGAGCAACAACACTTGTCTACGTTTGGGATACCTCTTTGGGAGAGGT	123420
Qy	1190	-----	1189
Db	123419	AAGCAATTAGTATTTGAGTAATGAGTGGTGAGAGTTGTTTCGCTCATGTAAAGTTTATG	123360
Qy	1190	-----AGTGTGTAATTGGATTTGGGATTCAGCAGGATTC	1224
Db	123359	ATTGAAACGTCAGGAGTTTTCTTCGAGGTTGTAATTGGATTTGGGATTCAGCAGGATTC	123300
Qy	1225	CACAGGAATAGAAATAAGCTGTTTGGGGTTGTGCAGCAGATGGAAGTGCANTTGTATACAGA	1284
Db	123299	CACAGGAATAGAAATAAGCTGTTTGGGGTTGTGCAGCAGATGGAAGTGCANTTGTATACAGA	123240
Qy	1285	AGTTGGGATTCAAATCTAA-----	1303
Db	123239	AGTTGGGATTCAAATCTAAAGGTTAGTACTTAGTAGTGTCTCTGCTTATCATTTTACCATG	123180
Qy	1304	-----AGATATCGG	1312
Db	123179	TTATGTTTGTAAAGAAATTAGAGGATAGATGAAATTTTCGTTGAAATTCACACAGATATGGG	123120
Qy	1313	CGTTTGGAGGACACAGGAGTGTATTTGAGAAGATTTTAAAGAAAGTAGGAGTCAACGTC	1372
Db	123119	CGTTTGGAGGACACAGGAGTGTATTTGAGAAGATTTTAAAG-AAAAGTAGAGTCAACGTC	123061
Qy	1373	TCCAGTTGTTGGTTAAATATTTCTGTAGTCCGGGAAGTAAGGTTCCGTTTGTGTGGAAGGTGT	1432
Db	123060	TCCAGTTGTTGGTTAAATATTTCTGTAGTCCGGGAAGTAAGGTTCCGTTTGTGTGGAAGGTGT	123001
Qy	1433	TTGGTTTGAATATGAGTGAGTGGTTAGAGAAGATTTAAACTTCCTCTTTTGTGTGCTCTTG	1492
Db	123000	TTGGTTTGAATATGAGTGAGTGGTTAGAGAAGATTTAAACTTCCTCTTTTGTGTGCTCTTG	122941
Qy	1493	ATTATTTATTTCTTTCATTTGGGAACATAACTCCCTTCAACAGCTACTCAATGTGAAATCT	1552
Db	122940	ATTATTTATTTCTTTCATTTGGGAACATAACTCCCTTCAACAGCTACTCAATGTGAAATCT	122881
Qy	1553	GTAATCAATTGTGTACCCACAGTCTTTTACTTTACTATCATCTCTTTCATATTGAAACGAG	1612
Db	122880	GTAATCAATTGTGTACCCACAGTCTTTTACTTTACTATCATCTCTTTCATATTGAAACGAG	122821
Qy	1613	AAGATAAAACGCTACTA	1629
Db	122820	ACGATAAAACGCTACTA	122804

RESULT 7
AF170921
LOCUS
DEFINITION
ACCESSION
VERSION

KEYWORDS	Pisum sativum (pea)	Db	471	TTGATGGATGGGAATCTTAATGTTTACGGATCTTAGTGACATAAAGGTTATGTTTC	530
SOURCE	Pisum sativum	Qy	687	GTGCTGTCAGTATCTCCAAATAGGATGCCACCTTATACACAGTTCCAGTGATCAAAAC	746
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.	Db	531	ATCTTGTGATGTTCCAGGTGAAGACACTCACTTAATCACTGTTTGGAGATCAGAC	590
REFERENCE	1 (bases 1 to 1526)	Qy	747	TTGATCTTATGGGATGTAACCTAGTCTCAAACTCTGTTTTTGGCGTGAATTCA	806
AUTHORS	Lapik, Y.R. and Kaufman, L.S.	Db	591	ATGTTGTTTATGGGATATTAATCTAGGCTTAGAACATCTGTTTGGAGCGAGTTTCA	650
JOURNAL	Unpublished	Qy	807	GTCTGGACATACCTGCTGATGTAAGCGTCTCAATCAGTGGATCAAAACCACTGGTT	866
REFERENCE	2 (bases 1 to 1526)	Db	651	GTCTGGACATACCTGCTGATGTAAGCGTCTCAATCAGTGGATCAAAACCACTGGTT	710
AUTHORS	Lapik, Y.R. and Kaufman, L.S.	Qy	867	TATATCTGTTTCATGCGGATCCAGACAGCGTGTGGGACACTGCTGTCGAAGCCGAGC	926
JOURNAL	Direct Submission	Db	711	TGATCTGTTCTTGGATGCGACTGCCAGATTTGGGACACTGCTGTTGGCAAGTCGAGC	770
FEATURES	Location/Qualifiers	Qy	927	AGTGGTACCTTTCATGCTCAGGAGGAGATGTTAATACGGTCAAGTTCTTCCGGATGG	986
source	1. 1526	Db	771	AGTGGGACATTTCAAGTATATATATCAGCAACACCAAGACCAAGATGTCACGTC	830
	/organism="Pisum sativum"	Qy	987	GTATAGATTTGGGACTGGATCAGACGATGGAACATGCAAGGCTGTATGACATAAGACTGG	1046
	/mol_type="mRNA"	Db	831	AAATAGATTTGGGACTGGATCAGACGATGGAACATGCAAGGCTGTATGACATAAGACTGG	890
	/cultivar="Alaska"	Qy	1047	TCACCACTCCAGTCT---ATCAGCCACATGTTGATGGTGGAGACGACCTGTCCACTC	1103
	/db_xref="taxon:3888"	Db	891	ACACCACTTCAAGTATATATATCAGCAACACCAAGACCAAGATGTCACGTC	950
	/tissue_type="apical bud"	Qy	1104	CATTCATCTCTGTCAGGAGACTCTTCTTCTGCTGGCTATGCGAGCAACCAACTTG	1163
	/dev_stage="seedling"	Db	951	CATTCATCTCTGTCAGGAGACTCTTCTTCTGCTGGCTATGCGAGCAACCAACTTG	1007
	/note="seedlings grown in 7 days of darkness or 5 days of darkness and 2 days of light"	Qy	1164	CTACGTTTGGGATACCTCTTCTTGGGAGAGTGTATTTGATTTGGATTCAGCAGGATTC	1223
	52..1185	Db	1008	CTATGTTTGGGATACCTCTTCTTGGGAGAGTGTATTTGATTTGGATTCAGGATTC	1067
CDS	/codon_start=1	Qy	1224	ACACAGGAATAGAATAAGCTGTTTGGGTTGTGACAGATGGAAGTGCATTGTGACAG	1283
	/product="G protein beta subunit"	Db	1068	TCATGAGGCGAGGATCACTCTGTTGGGTATGCTCTGATGGAAGTCTTTATGACAG	1127
	/protein_id="A049742.1"	Qy	1284	RAGTTGGGATCAAACTTAAGATATGGCGTTTGGAGGACACAGAGAGTCAATTGAAG	1343
	/db_xref="GI:5733806"	Db	1128	RAGTTGGGACACAAATTTAAAGATATGGGCAATTTGGAGGCGCATAGGAAGTGAATTGACT	1187
	/translation="MSVAELKHEHIAAETVNNLRERLQRRLSLDITDIAGYARQQ RAPVTFPTDILCRTHGKHYSLDTSKRNIVSAQDRLIVNALTYSQTHA IKLPCAWMTCAFPSTGQSVACGGLSDVCSIFNLNPTDRDGNLNVSMLSHGKYSV SCQYVPGDTHLITGSGDQCTVLWDITTLGLRTSVFSGFQSGADVLISISINGNSK LPSVSCDATALRWDTRVARSRAVTFHGEHGVNSKFPDPGNRFGTSGEDGTCRLFD INTGHLQYNOQHODNEMAHVTISAFISGRLLIAGTNGDCYVWDILLAKVNLIG SLQNSHEGRIITCLGMSADGSALCTGSWDTNLKIWAPFGHRRVI"	Qy	1344	RAGATTTAAAGAAAGTAGGAGTCAAGTCTCCAGTGTGTTGGTTAATATATATTCTGTAGTCG	1403
		Db	1188	CAATGTTAGGCTTCACTGTTTAAATGATGCTTGTGATATTGACTTTGATCCAGNAATTG	1247
ORIGIN		Qy	1404	GGAAGTAAGGTT 1415	
	Query Match 40.6%; Score 661.6; DB 15; Length 1526;	Db	1248	GAAGCAAGTT 1259	
	Best Local Similarity 72.8%; Pred. No. 5e-178;	RESULT 8			
	Matches 882; Conservative 0; Mismatches 324; Indels 6; Gaps 2;	LOCUS			
Qy	207 AATGTCGTCTCGAGCTCAAGAACGCCACGCCGTCGTACGAGACCGTTAATAACT 266	DEFINITION	NT284820	N.tabacum mRNA for G protein beta subunit (clone NcGbeta1).	
Db	51 AATGTCGTCTCGAGCTCAAGAACGCTCACATAGCAGCGACGGAACCGTTAACT 110	ACCESSION	Z84820		
Qy	267 CCGTACAGCTTAGACAGACGCGCTCCAGCTCCGATACCGATGTCGCGAGGTATTC 326	VERSION	Z84820.1	GI:1835160	
Db	111 CAGAGAACGATTGAAGCAGAGACGGCTTCTTGTGCTTGATACAGATATTGCTGGATATGC 170	KEYWORDS		beta subunit; g protein.	
Qy	327 AGCGGCGAAGGACGTAAGTCTCGGTGAGCTTCGAGCAAGGATCTGTTTGTGCTGAC 386	SOURCE		Nicotiana tabacum (common tobacco)	
Db	171 TAGGCTCAAGGTAGAGCTCTGTTACTTTTGTGCTCCACTGATATTCTTTGCTGTAGAAC 230	ORGANISM		Nicotiana tabacum	
Qy	387 TCTTCAGGACACACCGGAAAGGTTTATTATTAGATTTGACACCGAGAGAACCGGAT 446			Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.	
Db	231 GCTTCAAGGTATACCGGAAAGGTTATTATTATTGATTTGAGTTCAGAAAAGATAGGAT 290	REFERENCE		1 (bases 1 to 1430)	
Qy	447 TGTCAAGTATCAAGATGGGAGATTAATCGTGGGATGCTCTTAACAGTCAAGAAAC 506	AUTHORS		Lein, W. and Saalbach, G.	
Db	291 TGTAGTGCATCCAAAGATGGAAGATTAATAGTGTGGAAATGCTCTTAACAGCCAGAAAC 350	JOURNAL		Unpublished	
Qy	507 TCATGCTATTAACTCCCTTGTGTCATGGGTTATGACATGCTTCTCTCCAAATGGTCA 566	REFERENCE		2 (bases 1 to 1430)	
Db	351 TCAAGTATTAAGCTTCTTTGTCATGGGTATGAGTGTGCTTTCTACCAACTGGTCA 410				
Qy	567 GTCGGTGGGTGGTGGATTAAGACAGTGTATGTTCTTATCTTTAGCGCTTAGCTCAACGGC 626				
Db	411 ATCTGTTGCTTGTGGGGCTTGACAGTGTGTTGCTCTATTCTTCAATCTTAAATCTCCAC 470				
Qy	627 GGACAAAGGATGGAACGTGACCGGTTTCAAGAAATGCTCACTGCTCAGCGGATATGTTTC 686				

AUTHORS Lein, W.
Direct Submitted (04-FEB-1997) Lein W., Institut fuer Pflanzen-genetik und
TITLE Kulturpflanzenforschung, Molecular Cell Biology, Correnstrasse 3,
JOURNAL Gatersleben, Sachsen-Anhalt, Germany, D-06466
FEATURES Location/Qualifiers
source 1. .1430
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/strain="GR1"
/db_xref="taxon:4097"
/clone="NcObetal"
/tissue_type="leaf"
/clone_lib="cDNA library"
1. .72
73. .1206
/codon_start=1
/product="G protein beta subunit"
/protein_id="CAB06618.1"
/db_xref="GI:1835161"
/db_xref="GOA:P93397"
/db_xref="InterPro:IPR001632"
/db_xref="InterPro:IPR001680"
/db_xref="UniProt/Swiss-Prot:P93397"
/translaton="MSYTELKERHMAATQTVNDRLEKQKRLQLLDVSGYARSG
KTPVTFGPTDLVCRCRI LQGHGKVYSLDWTPEKNRIVASODGRLI VVNALTSOKTHA
IKLPCAWMTCAVSCPSVACGLDVSCTFINLSPIDKGNHPVSRMLSGHKYVPS
SCQVPEBDTHLITSSSDQTCVLWDITGLRTSVFGFOSGHTADVQSVISISSNP
LFVSGSDTARLMDTVRAQRKTFYHGEHDVNVKFPDGNRFGVSGSDGTCRLPD
IRTEHQIQVYVPHGDDIPIHVTSMAFSISGRLLFVGSNGDCVYVNDTLAKVYLING
GVQNSHEGRISCLGLSADGSALCTGSWDTNLKIWAFGGHRSVI"
1207. .1430
3'UTR 1414
polyA_site
ORIGIN
Query Match 40.3%; Score 656.4; DB 15; Length 1430;
Best Local Similarity 74.8%; Pred. No. 1.5e-176;
Matches 851; Conservative 0; Mismatches 281; Indels 6; Gaps 2;
Qy 207 AATGCTGTCTCCGAGCTCAAGAAGCCGACCGCTGCTACGAGACCGGTATTAACCT 266
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 72 AATGTCAGTGCACAGCTGGAAGAGCGGCATATGCGCGCTACACAGACTGTAAATGATCT 131
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 267 CCGTGACCGAGCTTAGACAGACAGCGCTCCAGCTCTCGATACCGATGTGGCGAGGTATTC 326
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 132 CCGTGAAAACCTTAAGCAGAGAGCTCTCCAAATTACTGCACACTGATGTTCTCGATATGC 191
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 327 AGCGGCGCAAGGACGTACTCGGTTGAGCTTCGAGACCAACGGATCTGGTTGTTGTCGTAC 386
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 192 AAGTTCGCAAGGTAAACCTCCGGTCACTTTGGCCCAACAGACTCTGGTTGTTGTTAGGAT 251
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 387 TCATTACGGGACACACCGGAAGGTTTATTCAATTAGATTGGACACCGGAGAGGAACCGGAT 446
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 252 CCTCAAGGACACACTGGAAGGTATATTCACTGGAATTGGACTCCAGAAAAGATTCGTAT 311
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 447 TGTCAAGTGCATCTCAAGATGGAGATAAATCGTGTGGAATGCTCTACGAGTTCAGAAAC 506
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 312 AGTCAGTGCATCCCAAGATGGCATTAAATAGTGTGGAATGCTCTCAAGCCAGAAAC 371
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 507 TCATGCTATTAAACTCCCTTGTGATGGGTTATGACATGTCTTTCTCTCCAAATGGTCA 566
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 372 CCATGCAATTAAAGTTCCGTTGCTTGGGTTATGACCTGCGCCTTCTCTCTAGTGGCA 431
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 567 GTCGGTTGCGTGGTGGAATTAGACAGTGTATGTTCTATCTTTAGCCTTAGCTCAACGGC 626
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 432 GTCTGTTGCGCTGCGGTGGCTTGAACAGTGTCTGCTCTATCTTCAACTTAAATTCGCCAAT 491
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 627 GGACAGGATGGAACGTGTACCGGTTTCAAGAAATGCTCACTGGTCACAGGGGATATGTTTC 686
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 492 CGATAAGATGGGAACCACTCTGTATCAAGAAATGCTTAGTGGGCATTAAGGGTTATGTGTC 551
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 687 GTGCTGTCAGTATGTCCCAATAGGAGTGCACCACCTTATCACCAAGTTCAAGTGTATCAAC 746
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

/db_xref="taxon:4097"
/tissue_type="hairy root"
143..1276
/note="NtGB1"
/codon_start=1
/product="heterotrimeric GTP-binding protein subunit beta
1"
/protein_id="AAG12330.1"
/db_xref="GI:10048265"
/translation="MSVKELKRRHMAATQTVNDLREKLKRLQLDLDVDSGYARSQG
KTPVIFGPTDLVCCRILQHGKVKVSLDMTPKRNIVASQDGRILVNNALTSOKTHA
IKLPCAWMTCAFPSGSGVACGLDVSCLIFNLSPDKGNHPVSMLSGHKGYVS
SCQVYDPTDLHITSQSDQTCVLWDITTLGLRTSVFGGFSQSHADVQSVSISSNPR
LFWSSCDITAGLWDRVARSRAQRTFYHGEQDVTYKSPGNRFGTSBGTCLRLFD
IRTHQLOVYQPHGDGDIPIHVTNSAFSISGRLLFVYSGNDGVYDWWLLAKVNLNG
GYVNSHGRISICGLSADGSALCTGSDWNTNLKIWAFGGHRSVI"

ORIGIN
Query Match 40.2%; Score 654.8; DB 15; Length 1600;
Best Local Similarity 74.7%; Pred. No. 4.5e-176;
Matches 850; Conservative 0; Mismatches 282; Indels 6; Gaps 2;

QY 207 AATGTCGTCTGCCAGCTCAAGAAGCGCCACGCCGTCGTACGGAGACCGTTAAATAACCT 266
DB 142 AATGTCAGTGA CAGAGCTCAAGAGCGGCATATGCGCGCTACACAGACTGTAATATGATCT 201

QY 267 CCGTGACCAAGCTTAGACAGACGCTCCAGCTCTCGATACCGATGCGAGGATATTC 326
DB 202 CCGTGAAAGAACTTAAGCAGAGCGTCTCAATTACTCGACACTGATGTTCTCGATATGCG 261

QY 327 ACGGGCGCAAGACGTACTCGGTGAGCTTCGGAGCAAGGATCTGGTTGTTGTCGTAC 386
DB 262 AAGGTGCGAAGTAAACCTCGGTCACTTTGGCCCAACAGATCTGGTTGTTGTTAGGAT 321

QY 387 TCTTCAGGACACACCGGAAAGTTTATTCATTAGATTGGACACCGAGAGAAACCGGAT 446
DB 322 CCGTAGAGACACCTGGAAGATATTAATCTGAGATTGGATCTCCAGAAAGATGCTAT 381

QY 447 TGTCAAGTGCATCTCAAGATGGAGATTAATCGTGTGGAATGCTCTTAACAGATCAGAAAC 506
DB 382 AGTCAGTGCATCCCAAGATGGCAGATTAATAGTGTGGAATGCTCTCAAGCCAGAAAC 441

QY 507 TCATGCTATTAACCTCCCTGTGTCATGCGGTATGACATGCTCTTCTCCAAATGGTCA 566
DB 442 CCATGCAATTAAGCTTCCTGTGCTTGGGTATGACCTCGCGCTTCTCTCTAGTGGGCA 501

QY 567 GTCGGTGTGCTGTGGGATGACAGCTGATGCTTCTATCTTTAGCTTTAGCTCAACGGC 626
DB 502 GTCGTGTGCTGCGGTGGCTTGACAGTGTCTGCTCTATCTTCAACTTAATTCGCCAAT 561

QY 627 GGACAAAGGATGGAACCTGTACCGGTTTCAAGAAATGCTCACTGGTTCACAGGGGATATGTTTC 686
DB 562 CGATAGGATGGGAACCATCTGTATCAAGAAATGCTTAGTGGGCAATAAGGTTATGTGTC 621

QY 687 GTGCTGTAGTATGTCCTCAAAAAGAGATGCCACCTTATCAGATTCACAGTTGAGTGATCAAC 746
DB 622 TTCTCTGCAATATGTTCCAGATGAGGATATCTCACTTAATAACTAGTTCTGGGTGATCAAC 681

QY 747 TTGTATCTATGGGATGAATCTACTGCTCAAAAACCTCTGTTTGGCGGTGAATTTCA 806
DB 682 ATGTGTCTTTGGGATATTAATCTACTGCTTAAGAACTTCTGTTTGGAGGTGAGTTTCA 741

QY 807 GTCTGGACATPACTGCTGATGTAATAAGCGTCTCAATCAGTGGATCAAAACCCAACTGGTT 866
DB 742 ATCCGGGCACACTGCAGATGTAACAAGTGTCTCAATTAGTTTCATCAAAACCCAGACTGTT 801

QY 867 TATATCTGGTTCTATGGATTCACAGCAAGTGTGTGGGACATCTGCTGTGCAAGCCGAGC 926
DB 802 TGTATCTGGGTCCTGTGTGACAACTGCTGAGCTGTGGGACACCCGAGTTGCTAGTGCAGC 861

QY 927 AGTGGCTGACCTTTTCATGGTCACAGGGAGATGTTAATAGCTCAAGTCTTTCCCGATGG 986
DB 862 TCACAGAACATTTTATGGTCACAGGGAGATGTTAATCTGTAAGATTTCTCCCTGATGG 921

```

```

QY 987 GTATAGATTGGGACTGGATCAGACGATGGACATGCGAGGCTGTATGACATAAGGACTGG 1046
DB 922 TAATAGATTGGAACTGGTTCAGAGGATGGAACTGCGAGATTAATTGACATTAGGACTGG 981

QY 1047 TCACCAACTCCAGGT---CTATCAGCCACATGGTGTGATGGTGGAGAACGACCTGTCACTC 1103
DB 982 ACACCAAGCTCAAGTGTACTACACAGCGCATGGTGTGATGATATCCCTCATGTGACTTC 1041

QY 1104 CATTGCAATCTCTGTGTCAGGAGACATCTTTTTCGTCGGCTATCGGAGCAACACACTTG 1163
DB 1042 CATGCAATTTTCTATCTCAGGCCGCTTCTCTCTTTGTCGGATACTC---AAATGGTGATTG 1098

QY 1164 CTAGTTTGGGATACCTCTCTGGGAGAGGTGATGATGATTTGGATTTGAGATTCACAGCAGATT 1223
DB 1099 TTATGTGTGGGACACCTTATTAGCAAGGTGGTCTTAACCTTTGGAGGAGTTCAAACTC 1158

QY 1224 ACACAGGAATAAGATTAAGCTGTTTGGGGTTCCTCAGCAGATGGAAGTGCATTTGTGTACAGG 1283
DB 1159 TCATGAGGGCGAATAGTTCCTGGGACTGTGAGCTGATGAGAGCGCTTATGTGTACAGG 1218

QY 1284 AAGTTGGGATTCAAATCTTAAAGATATGGCGCTTTGGAGGACACAGGAGAGTATTGA 1341
DB 1219 AAGTTGGGATACAAACCTGAAGATTTGGGCTTTTGGAGGACACAGAGTGTGATCTGA 1276

RESULT 10
NTZ84821
LOCUS N.tabacum mRNA for G protein beta subunit (clone NtGbeta2).
DEFINITION N.tabacum mRNA for G protein beta subunit (clone NtGbeta2).
ACCESSION NTZ84821
VERSION 284821.1 GI:1835162
KEYWORDS beta subunit; g protein.
SOURCE Nicotiana tabacum (common tobacco)
ORGANISM Nicotiana tabacum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
REFERENCE 1 (bases 1 to 1434)
AUTHORS Lein, W. and Saalbach, G.
TITLE Characterization of G-protein function in plants
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1434)
AUTHORS Lein, W.
TITLE Direct Submission
JOURNAL Submitted (04-FEB-1997) Lein W., Institut fuer Pflanzen-genetik und
Kulturpflanzenforschung, Molecular Cell Biology, Corrensstrasse 3,
Gatersleben, Sachsen-Anhalt, Germany, D-06466
FEATURES
Location/Qualifiers
1..1434
/organism="Nicotiana tabacum"
/mol_type="mRNA"
/strains="SRI"
/db_xref="taxon:4097"
/clone="NtGbeta2"
/tissue_type="leaf"
/clone_lib="cDNA library"
1..74
75..1208
/codon_start=1
/product="G protein beta subunit"
/protein_id="NtGbeta2"
/db_xref="GI:1835163"
/db_xref="GOA:P93398"
/db_xref="UniProt/Swiss-Prot:P93398"
/translation="MSVKELKRRHMAATQTVNDLREKLKRLQLDLDVDSGYARSQG
KTPVIFGPTDLVCCRILQHGKVKVSLDMTPKRNIVASQDGRILVNNALTSOKTHA
IKLPCAWMTCAFPSGSGVACGLDVSCLIFNLSPDKGNHPVSMLSGHKGYVS
SCQVYDPTDLHITSQSDQTCVLWDITTLGLRTSVFGGFSQSHADVQSVSISSNPR
LFWSSCDITAGLWDRVARSRAQRTFYHGEQDVTYKSPGNRFGTSBGTCLRLFD
IRTHQLOVYQPHGDGDIPIHVTNSAFSISGRLLFVYSGNDGVYDWWLLAKVNLNG
GYVNSHGRISICGLSADGSALCTGSDWNTNLKIWAFGGHRSVI"
1209..1434
5'UTR
CDS
3'UTR

```

polyA_site 1416		ORIGIN	
Query Match 40.0%; Score 651; DB 15; Length 1434; Best Local Similarity 74.5%; Pred. No. 5.5e-175; Matches 848; Conservative 0; Mismatches 285; Indels 6; Gaps 2;			
QY	207	AATGTCCTCTCCGAGCTCAAGAAAGCGCAGCCGTCGCTACGAGACCGTGAATAAACCCT	266
DB	74	AATGTCAGTGAAGAGCTGAAGAGCGCATATGCGCGCTACACAACTGTAAATGATCT	133
QY	267	CGGTGACAGCTTAGACAGAGCGCTCCAGCTCTCGATACCGATGCGGAGGATATTC	326
DB	134	CGGTGAAGAACTTAAGCAGAGCGCTCTCAATTACTCGACACTGATGATCTCGGGATATGC	193
QY	327	AGCGCGCAAGAGCGTACTCGGGTGAAGCTTCGGAGCAACCGATCTGGTTGTGTCGTAC	386
DB	194	AAGGTGCGAAGTAAACTCCGGTCACTCTTGGCCCAACAGATCTGGTTGTGTAGGAT	253
QY	387	TCTTCAGGACACACCGGAAGGTTTATTCATTAGATTGGACACCGGAGAGAACCGGAT	446
DB	254	CCTGCAAGGACACACTGGAAAGGTATATTCATCTGGATTGGACTCCAGAAAAGAAATCGTAT	313
QY	447	TGTCAGTGCATCTCAAGATGGAGATTAAATCGTGTGGAATGCTTAACGAGTCAAGAAAC	506
DB	314	AGTCAGTGCATCCCAAGATGGCAGATTAAATAGTGTGGAAATGCTCTCAACGCCAGAAAC	373
QY	507	TCATGCTATTAAACTCCCTTGTGCAATGGGTTATGACATGTGCTTTCTCTCCAAATGCTCA	566
DB	374	CCATGCAATTAAAGCTTCATGCTTGGTTATGACCTGCGCTTCTCTCTAGTGGCA	433
QY	567	GTCGGTTCGGTGTGTGATAGACAGTGTATGTTCTATCTTTAGCCCTTAGCTCAACGGC	626
DB	434	GTCGTGTCCTCGCGTGGCCCTTGACAGTGTCTGCTCTATCTTCAACTTAAATTCACCGAT	493
QY	627	GGACAAGGATGMACTGTACCGTTTCAAGATGCTCACTGCTCAACAGGGATATGTTTC	686
DB	494	CGATAGAGATGGAAACCACTCTGTATCAAGAAATGCTTAGTGGCATAGAGGATATGTTTC	553
QY	687	GTGCTGTCAAGTATGTCACAAATAGGATGCCACCTTATCAACAGTTCAGGTGATCAAC	746
DB	554	TTCCCTGTCAAGTATGTCACAGATGAGATACACGTAATAACTAGTCTCGGTGATCAAC	613
QY	747	TGTATCTTATGGGATGTAACTAGTGTCTCAAACTTCTGTTTGGCGGTGAATTTCA	806
DB	614	ATGTGTCTCTTGGGATATACTACTGGCTTAAGAACTTCTGTCTTGGAGGTGAGTTTCA	673
QY	807	GTCTGGACATCTGCTGATGTACTAAGGTCTCAATCAGTGTGATCAAAACCCAAACTGGT	866
DB	674	ATCCGGGCACACCGCAGATGTACAAAGTGTCTCAATTAAGTTCATCAAAACCCCAAGACTGT	733
QY	867	TATATCTGGTTCATGCGATTCCACAGCAGTGTGGGACACTCTGCTGCAAGCCGAGC	926
DB	734	TGTGTCTGGTCTGTGACTCAACTGCTCGACTATGGGACACCCGAGTGTGCTAGTCGAGC	793
QY	927	AGTGGTACCTTTCAATGTCACAGGAGATGTTAAACGGTCAAGTCTTCTTCCGATGG	986
DB	794	TCAACGAAACATTTTATGCTCATGAGGAGATGTTAACTACTGTAAAGTCTTCCCTGATGG	853
QY	987	GTATAGATTTCGGACTGATCAGAGATGGAACATGCAAGCTGTATGATCAATAGGACTGG	1046
DB	854	TAATAGATTTCGAACCTGTTTCAATGATGGAACCTGCGAGATTAATTTGATCAATAGGACTGG	913
QY	1047	TCACCAACTCCAGGT---CTATACGCCACATGGTGTGATGGTGGAGAACCGACCTGTCACTC	1103
DB	914	ACACCAGCTGCAAGGTACTACCGCCGATGGTGTGATGATATCCCTCATGTGACTTC	973
QY	1104	CATTGCATCTCTGTGTACGGGAGACTTCTTTTCGTGGCTATGCGGACAAACACTTTG	1163
DB	974	CATGGCATTTTCTATCTCAGGCCGCTTCTCTCTTTGTCGGGTACTC---AAATGGTGTATG	1030
QY	1164	CTACGTTTGGGATACCTCTTTGGGAGAGGTGTATTTGGATTGGGATTTACAGGAGATTC	1223
Db	1031	TTATGTGTGGACACCCCTATTAGCAAGGTGGTCTCTAAACTTTGGAGCAGTTCAAAACCTC	1090
QY	1224	ACACAGGAATAGAAATAGCTGTTTGGGGTTTCTACGACAGATGGAAGTGCATTGTGTACAGG	1283
DB	1091	TCATGAAGGGCGAATAAGTTTGCCTGGGACTGTCTAGCTGATGGGAGCGCTTATGTACAGG	1150
QY	1284	AAGTTGGGATCAAAATCTAAAGATATGGCGTTTGGAGGACACAGGAGAGTATTGAA	1342
DB	1151	AAGTTGGGATACAAACCTGAAGATTTGGGCTTTTGGAGGGCACAGAAGTGTGTATCTGAA	1209
RESULT 11			
NPY09513			
LOCUS	NPY09513	1505 bp	linear
DEFINITION	N.plumbaginifolia mRNA for G protein beta subunit.		
ACCESSION	Y09513		PLN 04-MAY-2000
VERSION	Y09513.1	GI:1695178	
KEYWORDS	G protein beta subunit.		
SOURCE	Nicotiana plumbaginifolia (curled-leaved tobacco)		
ORGANISM	Nicotiana plumbaginifolia		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiids; Solanales; Solanaceae; Nicotiana.		
AUTHORS	1 Kaydamov, C., Tewes, A., Adler, K. and Manteuffel, R.		
TITLE	Molecular characterization of cDNAs encoding G protein alpha and beta subunits and study of their temporal and spatial expression patterns in Nicotiana plumbaginifolia Viv		
JOURNAL	Biochim. Biophys. Acta 1491 (1-3), 143-160 (2000)		
PUBMED	10760577		
REFERENCE	2 (bases 1 to 1505)		
AUTHORS	Kaydamov, C.S.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-NOV-1996) C.S. Kaydamov, Institute of Plant Genetics, Molecular Genetics, Serology Group, Correnestrae 3, Sachsen-Anhalt, Gatersleben, 06466, FRG		
FEATURES	Location/Qualifiers		
source	1. 1505		
	/organism="Nicotiana plumbaginifolia"		
	/mol_type="mRNA"		
	/db_xref="taxon:4092"		
	/clone="pNPGPBET"		
	/tissue_type="somatic embryos derived from mesophyll protoplasts"		
	/clone_lib="lambda ZAP Express"		
	124. 1257		
	/codon_start=1		
	/product="G protein beta subunit"		
	/protein_id="CAA70704.1"		
	/db_xref="GI:1695179"		
	/db_xref="GOA:P93339"		
	/db_xref="UniProt/Swiss-Prot:P93339"		
	/translation="MSVTEIKERHMAOTVNDLREKLKOKRLQLLTDVSGYARSG KTPVTPGDLVCCRILQHTGKVSLDWTFPEKNRIVSASQDGRLLIWNALTSKTHA IKLPAAVMTCFSPSGVACGLDSCSIFNLNPEIDKDNHVPRLMSLGHGYVS SCQYVPDEDTHLITSSGDQTCVLMDITTLGLRTSVFGFQFGHTADTVQSVSISNRP LFVSGSDTTLARLMDTRVASRAQRTFYCHEGDVNTVFPDGNRFGTSGEDGTCRLFD IRTGHQIQVYYPQHGDDIPIHTSMAPSISGRLLFVRSNGDCVMDTLAKVVLNLG AVQNSHSGXISCLGLSADGXXLCTGSDWTNLKIWAFGGHRSVI"		
CDS			
ORIGIN			
Query Match 39.6%; Score 644.8; DB 15; Length 1505; Best Local Similarity 74.0%; Pred. No. 3.3e-173; Matches 843; Conservative 1; Mismatches 289; Indels 6; Gaps 2;			
QY	207	AATGTCCTCTCCGAGCTCAAGAAAGCGCAGCCGTCGCTACGAGACCGTGAATAACCT	266
DB	123	AATGTCAGTGACAGAGCTGAAGAGCGGCATATGCGCGCTACACAGACTGTAAATGATCT	182
QY	267	CCGTGACAGCTTAGACAGAGACCGCTCCAGCTCTCCGATACCGATGTGGCAGGATTC	326
DB	183	CCGTGAAAAACTTAAGCAAAAGCGTCTCCAACTACTCGACACTGATGTTCTTGTGATATGC	242

QY 327 AGCGCGCAAGGACGTACTCGGTGAGCTTCGGAGCAACGGATCTGGTTTGTGTCTGCTAC 386
DB 243 AAGGTGCAAGGTAATAAATCTCGGTGACCTTTGGCCCAACAGATCTGTGTGTGTAGGAT 302
QY 387 TCTTCAGGACACACCGGAAAGTTTATTCATTAGATTGGACACCGGAGAGAACCGGAT 446
DB 303 CTGTCAGGACACACTGGAAAGGTATATCTAGATTGGACTCAGAAAGATCTGAT 362
QY 447 TGTCAAGTCACTCAAGATGGAGATTAATCGTGTGGAATGCTTAAACAGGTCAAGAAAC 506
DB 363 AGTCAGTGCATCCCAAGATGGCAGATTAATAGTGTGGAATGCTCTCAACAGCCAGAAAC 422
QY 507 TCATGCTATTAACTCCCTTGTGCATGGTTATGACATGTGCTTTCTCTCCAAATGTCAC 566
DB 423 CCATGCAATTAAGCTTCCATGTGCTTGGGTTATGACCTGCGCTTCTCTCTAGTGGCA 482
QY 567 GTCGGTGGTGTGGTGTAGACAGTGTATGTTCTATCTTTAGCCTTAGCTCAACGCG 626
DB 483 CTCGTGTGCTGTGGCTTGAACAGTGTGCTCTATCTTCACTTAATTCACCAAT 542
QY 627 GGACAAAGGATGGAATGTACCGGTTTCAAGAAATGCTCACTGTGTCAAGGGATATGTTTC 686
DB 543 CGATAAGGATGGCAACCATCTGTATCAAGATGCTTAGTGGGCATTAAGGGTTATGTGTC 602
QY 687 GTGCTGTAGTATGTCCTCAAAATGAGATGCCACCTTATCACCAGTTCAGGTGATCAAC 746
DB 603 TTTCTGTCAATGTTTCCAGATGAGGATACTCACCTTAATACAGTTCCTGCTGATCAAC 662
QY 747 TTGTATCTTATGAGATGTAATCTACTGCTCAAACTTCTGTTTGGGGTGAATTTCA 806
DB 663 ATGTGCTCTTGGGATATTAACCTACCGTCTAAGAACTTCTGCTTGGAGGTGAGTTCA 722
QY 807 GTCTGACATCTGCTGATGTACTAAGCTCTCAATCAGTGGATCAAAACCCAACTGGTT 866
DB 723 ATTCGGGCACACCGCAGATGTACAAAGTCTCAATTAGTTTCATCAAAACCCAGACTGTT 782
QY 867 TATATCTGTTTCAATGAGTTCACAGCAGGTGTGGGACACTCTGCTGTGCAAGCGAGC 926
DB 783 TGTATCTGGGCTCTGTGACACACGCTGCTGAGTGTGGGACACCGGAGTTGCTAGTCGAGC 842
QY 927 AGTGGGTACTTTCATGCTCAGGAGGATGTATACGCTCAAGTCTTCTTCCGATGG 986
DB 843 TCACGCAATTTTATGTTCAGAGGAGATGTATATCTGTAAAGTTCTTCCCTGACGG 902
QY 987 GTATAGATTGGGACTGTGATCAGACATGGAACATGCAAGCTGTATGATCAAGGACTGG 1046
DB 903 TAATAGATTGGAACTGGTTTCAGAGGATGGAACCTGCAGATTTATTTGACATTTAGGACTGG 962
QY 1047 TCACCAATCTCAGGT---CTATCAGCCACATGTTGATGTGAGAACGGACCTGTACCTC 1103
DB 963 ACACCAAGTGCAGTGTACTTACCGCCGATGTTGATGTTGATGATCCCTCATGTGACTTC 1022
QY 1104 CATTCGATCTCTGTCTCAGGAGACTTCTTTTCGCTGCTATGCGAGCAACAACTTG 1163
DB 1023 CATGGCATTTCTATCTCAGGCGCTTCTCTTTTGGCGGTACTC---AAATGGTGATG 1079
QY 1164 CTACGTTTGGGATACCTCTTGGGAGAGTGTATTTGGATTTGGGATTTACAGCAGGATTC 1223
DB 1080 TTATGTGGGACACCTATTAGCAAGGTGGTCTTAACCTTGGGAGGAGTTCAAACTC 1139
QY 1224 ACACGAAATAGAAATAGCTGTTTGGGGTGTGACAGATGGAAGTGCATTTGTGACAGG 1283
DB 1140 TCATGAAGSSGGATAAGTTGCTTGGACTGTGAGTGTGGGAGSCCTTATGTACAGG 1199
QY 1284 AGCTCGGATTCAAATCTAAAGATATGGGCTTTGGAGGACACAGGAGAGTGAATTTGAA 1342
DB 1200 AGTTGGGATACAAACCTGAAGATTTGGGCTTTTGGAGGACACAGAAAGTGTGATCTGAA 1258

RESULT 12

AF414114

LOCUS

DEFINITION Solanum tuberosum G protein beta subunit 2 (GB2) mRNA, complete

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISMAF414114
AF414114.1 GI:15778631

Solanum tuberosum (potato)

Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum.REFERENCE
AUTHORS
TITLE1 (bases 1 to 1540)
Kang, S.G., Lee, H.J., Park, E.H. and Suh, S.G.
Molecular cloning and characterization of cDNAs encoding
heterotrimeric G protein alpha and beta subunits from potato
(Solanum tuberosum L.)
(Solanum tuberosum L.)
Mol. Cells 13 (1), 99-106 (2002)JOURNAL
PUBMED

11911481

REFERENCE

2 (bases 1 to 1540)

Kang, S.-G. and Lee, H.-J.

AUTHORS

Direct Submission

TITLE

Submitted (27-AUG-2001) Institute of Biotechnology, Yeungnam

JOURNAL

University, Kyongsan 712-749, Kyongsan 712-749, Korea

FEATURES

source

Location/Qualifiers

1..1540

/organism="Solanum tuberosum"

/mol_type="mRNA"

/db_xref="taxon:4113"

1..1540

/gene="GB2"

/gene="GB2"

/codon_start=1

/product="G protein beta subunit 2"

/protein_id="AAL07488.1"

/db_xref="GI:15778632"

/translation="MSVAELKERHMAATQTVNDLREKLQKRLQLDLDVSGYAKTQ
KLPVFTGDLVCRILQGHGKVVSLDWTPEKNRIVSASQDGRILVNNALTSQKTHA
IKLPCAVMTCAFPSPGSGSVAGGLDSACSIPNLNSPIDKGIHPVSRMLSGKYVS
SCQVPDTHLITSSGDQTCVLWDITGLRTSVFGFQSGHTADSVSLSSNPK
LFGSGCDTARLMDTRVRAQRTFHGHSDVITVFPDGNRGTGDSGDSCLRFD
IRTGQLQVNPQPHGDGDIPIHTSIATFISGIRLLFVGSNGDVCYWDTLAKVNLG
SVQNSHEGRISCLGLSDGSALCTGSDMTNLKIWAFGGHRSVI"

ORIGIN

Query Match 39.4%; Score 642.4; DB 15; Length 1540;
Best Local Similarity 73.9%; Pred. No. 1.6e-172;
Matches 843; Conservative 0; Mismatches 291; Indels 6; Gaps 2;

QY 205 GGAATGTCTGTCGAGCTCAAGAACCCACGCGCTCGCTACGAGACCGCTTAATAAC 264
DB 93 GAAATGTCAATGTCGAGCTCAAGAACCGGACACATGCGCGCTACACAGACTGTAATGAT 152
QY 265 CTCCTGACACAGCTTAGACAGAGACCGCTCCAGCTCTCGATACCGATGTGCGAGGTAT 324
DB 153 CTCCTGAAAACTTAAGCAGAAGCGTCTCCAATTAATCTACGACACAGATGTTCTCGGTAT 212
QY 325 TCAGCGGCGCAAGACGTAATCTCGGTGAGCTTCGGAGCAACCGATCTGTTGTGTCT 384
DB 213 GCAAGACGCAAGGTAATACTCCGTAACGTTTCGCGCCCAACAGATCTAGTTTGTGTAGG 272
QY 385 ACTCTTCAGGGACACACCGAAAGGTTTATTCTATTAGATTGGACACCGGAGGAGAACCGG 444
DB 273 ATCTGCAAGGACACATGGAAGGCTTATTCTACCTGAGCTGAGCTCTCTGAAAAATCTGT 332
QY 445 ATTGTCAAGTCACTCAAGATGGAGATTAATCTGTTGGAATGCTCTAAACGAGTCAGAAA 504
DB 333 ATAGTCAGTGCATCCCAAGATGTAATTAATAGTGTGGAATGCTCTCAAGCCAGAAA 392
QY 505 ACTCATGCTATTAAATCTCCCTTGTGCAATGGGTATGACATGTGCTTCTCTCAATGTT 564
DB 393 ACCCATGCAATTAAGCTTCCATGTGCTGGGTATGACCTGTGCTCTCTCTCTAGTGA 452
QY 565 CAGTCGTTGCTGCTGCTGATGATGATGCTTCTATCTTTAGCTTCTAGCTCAACG 624

Db 453 CAGTCTGTTGCTTGTGGCGGCTTGACAGTGCCTCTATCTTCAACTTAATTCACCA 512
QY 625 GCGGACAAGAGTGAACCTGTACCGTTTCAAGAAATGCTCACTGTGTCA CAGGGGATATGTT 684
Db 513 ATTGATAAGGATGGGATCCATCCAGTATCGAGATGCTTAGTGGCATAAGGGGTATGTG 572
QY 685 TCGTGTCTCAGTATGTCCTCAAAATCAGGATGCCACCTTATCACCAGTTCAGGTGATCAA 744
Db 573 TCTTCGTGTGAGTATGTTCCGGATGAGGATATCTACCTTAATAACTAGTATCTGGTGATCAA 632
QY 745 ACTTGATCTTATGGGATGTAATCTGCTTCAAAAATCTTGTGTTTGGGGGTGAATTT 804
Db 633 ACATGTGTACTTTGGGATATAACTACTGCGCTTAAGAACTTCTGTGTTTGGAGGTGAGTTT 692
QY 805 CAGTCTGACATATCTGCTGATGTAATAAGCGTCTCAATCAGTGGATCAAAACCAACTGG 864
Db 693 CAATCTGGGCACATCTGACAGATGATCAAGTGTCTCAATTAGTTTATCTTAACCCCAACTA 752
QY 865 TTTATATCTGTTTATCGATTCACAGCATCCACAGCAGGTTGTGGGACACTCGTGTGCAAGCCGA 924
Db 753 TTTGTGTCTGGGTCTGTGACACAACTCTCGACTGTGGGACACCCGAGTTGCTAGTCTGA 812
QY 925 GCAGTGCCTACTCTTTCATGGTTCACGAGGAGATGTTAATACGGTCAAGTCTTTTCCGGAT 984
Db 813 GCTCAACGAACATTTTCATGGACACGAGAGTGATGTTACTCTGTAAAGTCTTCTCCTGAC 872
QY 985 GGGTATAGATTTGGGACTGGATTCAGACGATGAAACATGACGCTGTATGACATAAGGACT 1044
Db 873 GGTATAGATTTGGAACCTGTTTCAGATGATGCGAGTGCAGATATTTGACATTAAGGACT 932
QY 1045 GGTACCAACCTCCAGGT---CTATCAGCCACATGCTGATGTTGAGAACGGACCTGTCAAC 1101
Db 933 GAACACCACTGCAAGTATACAAACACCGCATGTTGACGCTGACATCCCTCATGTGACT 992
QY 1102 TCCATTGCAATCTCTGTGTGAGGAGACTTCTTTTTCGCTGCTATGCGAGCAACAACT 1161
Db 993 TCCATTGCAATTTCTATCTCAGCGCTCTTCTTGTGCGGTACTC---TAATGGTGTAT 1049
QY 1162 TGCTACGTTTGGGATACCTCTTGGGAGAGGTTGATTTGGATTTGGGATTAACAGCAGAT 1221
Db 1050 TGTACGTTGGGACACCTTATTAGCAAGGTGGTCTTAACCTTAGGATCAGTTCAAAAC 1109
QY 1222 TCACACAGGAATAGATTAAGCTGTTTGGGTTGTCAGCAGATGGAAGTGCATTGTGTACA 1281
Db 1110 TCTCATGAAGCGCAATAGTGTCTGGACGTCTGAGCTGATGGAAGTGCCTTATGTACA 1169
QY 1282 GGAAGTTGGGATTTCAAATCTAAAGATATGGCGTTTGGAGGACACAGGAGTGAATTTGA 1341
Db 1170 GGAAGTTGGGATACAAACCTGAAGATTTGGGCTTTTGGAGGACACAGAAAGTGTGATCTGA 1229

RESULT 13
NTGPRBTBET 1560 bp mRNA linear PLN 18-APR-2005
LOCUS N.tabacum mRNA for G protein, beta subunit.
DEFINITION X98161
ACCESSION X98161.1 GI:1360091
VERSION cytoplasmic; G-protein beta subunit; signal transduction.
KEYWORDS Nicotiana tabacum (common tobacco)
SOURCE Nicotiana tabacum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Nicotiana.
1
Kunetssov.V.V. and Oelmuehler.R.
Isolation and characterization of cDNAs encoding the subunit beta
of heterotrimeric G proteins from N. tabacum (Accession No. X98161)
(PGR96-048)
JOURNAL Plant Physiol. 111, 948-948 (1996)
AUTHORS 2 (bases 1 to 1560)
Oelmuehler,R.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-1996) R. Oelmuehler, Botanisches Institut der

Ludwig-Maximilians-Universitaet, Menzingerstr 67, 80638 Muenchen,
FRG
Ref [3]: Plant Gene Register PGR96-048 (1996).
Location/Qualifiers
1. 1560
/organism="Nicotiana tabacum"
/molecule="mRNA"
/cultivar="Samsun NN"
/db_xref="taxon:4097"
/clone="5"
/tissue type="leaf"
/clone_lib="lambda gt11"
/dev_stage="seedlings"
120..1247
/function="signal transduction"
/note="beta subunit
localized in cytoplasm"
/codon_start=1
/product="G protein"
/protein_id="CAA66842.1"
/db_xref="GI:1360092"
/db_xref="GOA:Q40507"
/db_xref="InterPro:IPR001632"
/db_xref="InterPro:IPR001680"
/db_xref="InterPro:IPR011046"
/db_xref="UniProt/Swiss-Prot:Q40507"
/translation="MSVTEIKERHMAATQVSDLEKLRKQKRLQLLDTVDVSGYARSQ
KTPVFTGDTLCCRIQLQHTGKYSIDWTPEKRIIVASQDGLIIVNALSQKTHA
IKLPCAMVTCAPSGQSVACGLDSTYINLSPIDKDNHPVMSLGHKGYVS
SCQVPDEDTHLITSGSDQTCVLWDITTGRTSVFEGFOSGHTADVQSVISSNPR
LFTVSGSDTTARLWDRVARSARQTFYHGEDVNTVFPDGNRPFCTGSEDTGRLPD
IRTHQVQYIPHGHDGDIPIHVTSMFASISGRLLLVGVSYNGDCYVMDTLAKVVLNLG
GVQNSHGRISCLGLSADGSALCTGSWDTNLKIWAFGCTEV"

ORIGIN

Query Match 39.4%; Score 641.2; DB 15; Length 1560;
Best Local Similarity 74.5%; Pred. No. 3.6e-12; Mismatches 283; Indels 7; Gaps 3;
Matches 848; Conservative 0;
QY 207 AATGTCTGTCTCCGAGCTCAAAGAACCGCCGCTGCTACGAGAGCCGTTAATAACCT 266
Db 119 AATGTCACTGACAGAGCTGAAAGAGCGCATATGCGGCTACACAGACTGTAAGTGATCT 178
QY 267 CCGTGACAGCTTACAGACAGACGCCCTCCAGCTCTCGATACCGATGTGGCAGGATATTC 326
Db 179 CCGTGAAGAACTTAAGCAGAGAGCGCTCCCAATTAACGACACTGATGTTCTGGATATGC 238
QY 327 AGCGGCGCAAGGACGTACTCGGTGAGCTTCGGAGCAACGGATCTGTTCTGTCGTAC 386
Db 239 AAGGTCCGAGGTAAACCTCGGTGACCTTTGGCCCCAACAGATCTGGTTTGTGTAGGAT 298
QY 387 TCTTCAGGACACACCGGAAAGTTTATTCAATTAGATTGGACACCGGAGAGAACCGGAT 446
Db 299 CCGTCAAGGACACACTGGAAGGTATATTCACTGGATTGGACTCCAGAAAGATATCGTAT 358
QY 447 TGTCACTGATCTCAAGATGGGAGATTAATCGTGTGGAATGCTCTTAAGAGTTCAGAAAC 506
Db 359 AGTCAGTGCATCCCAAGATGGCAGATTAATAGTGTGGAATGCTCTCAAGCCAGGAAAC 418
QY 507 TCATGCTATTAAATCCCTTGTGCATGGTGTATACATGCTCTCTCCAAATGGTCA 566
Db 419 CCATGCAATTAAGCTTCGGTGTGCTCGGTATGACCTGCGCCCTCTCTCTAGTGGGCA 478
QY 567 GTCGGTTGCGTGTGGGATTAGACAGTGTATGTTCTATCTTATAGCCCTTACCTCAACGCG 626
Db 479 GTCTGTTGCGCTGCGGTGGCCCTTGACAGTGTCTGCTCTATCTACAACTTAATTCGCCAAT 538
QY 627 GGACAGAGTGAACCTGACCGGTTTCAAGATGCTCAAGATGCTCACTGCTCAGAGGATATGTTTC 686
Db 539 CGATAAGGATGGGAACCATCTCTGTATCAAGAAATGCTTAGTGGGATTAAGGGTTATGTGTC 598
QY 687 GTGCTGTCACTATCTCCCAATAGGATGCCACCTTATCACCAGTTCAGGTGATCAAAAC 746

```
Db      599 TTCCTGTCAATATGTTCCAGATGAGGATACTACCTTAATAACTAGTTCGTGTGATCAAC 658
Qy      747 TTGTATCTTATGGAGTAACTACTGCTCAAAACTTCTGTTTTTGGCGGTGAATTCA 806
Db      659 ATGTGTCTCTTTGGGATATAACTACTGCTCTAAGAACTTCTGCTTTGGAGGTGAGTTTCA 718
Qy      807 GTCCTGACATACCTGCTGATGTAAGCTCTCAATCAGTGGATCAAAACCCAACTGGTT 866
Db      719 ATCCGGGACACTGCAGATGTACAAGTGTCTCAATTAGTTTCAACAACCCAGACTGTT 778
Qy      867 TATATCTGTTTCATGCGATTCACAGCAGCGTTGTGGGACACTCGTGTCTGCAAGCCGAGC 926
Db      779 TGTATCTGGTCTCTGTGACAACTGCTCGACTGTGGGACACCGAGTTGCTAGTCGAGC 838
Qy      927 AGTGCGTACTTTTCATGCTGTCAGAGGAGATGTTAATACGGTCAAGTTCCTTCCGGATGG 986
Db      839 TCAACGAACTTTATGTTGTCAGAGGAGATGTTAATACCTGTAAGTTCCTCCCTGATGG 898
Qy      987 GTATAGATTGGGACTGGATCAGACATGGAACATGCAGCGCTGTATGATCAATAGGACTGG 1046
Db      899 TAATAGATTGGAACCTGGTTACAGAGATGGAACCTGCAGATTATTTGACATTAGGACTGG 958
Qy      1047 TCACCAACTCCAGT---CTATCAGCACAATGCTGATGCTGAGAACCGACCTGTCACTTC 1103
Db      959 ACACAGCTGCAAGTGTACTACACGCCGATGGTGTATTCCTCATGTGACTTC 1018
Qy      1104 CATTCGATCTCTGTCTCAGGAGACTTCTTTTCGCTGCTATGCGAGCAACAACTTG 1163
Db      1019 CATGGCAATTTCTATCTCAGGCGCTCTTCTCTTTGTCGGATATC---AAATGGTGATTG 1075
Qy      1164 CTACGTTTGGGATACCTCTGGGAGAGTGTATTGGATTGGGATTACAGCAGATTC 1223
Db      1076 TTATGTGTGGGACACCTATTAGCAAGGTGTCCTAACTTTGGGAGGAGTTCAAAACCTC 1135
Qy      1224 ACACAGGAATAGATAAGCTGTTTGGGGTTGTGAGCAGATGGAAGTGCATTGTGTACAGG 1283
Db      1136 TCATGAAGGGCGATTAAGTTGCTTGGGACTGTGAGTGTATGAGGCGCTTTATGTACAGG 1195
Qy      1284 AAGTTGGGATTCAAATCTAAGATATGGCGTTTGGAGGACACAGAGAGTGAATTGA 1341
Db      1196 AAGTTGGGATCAAACTGAAGATTTGGGCTTTGGAGG-GACAGAAGTGTGATCTGA 1252
```

RESULT 14

```
STGB1GENE      1524 bp      mRNA      linear      PLN 18-APR-2005
LOCUS          S.tuberosum mRNA for beta 1 subunit of heterotrimeric G-protein.
DEFINITION     X87837
ACCESSION      X87837.1 GI:1771733
VERSION        beta-1 subunit; g protein; gb1 gene; GB1 protein.
KEYWORDS       Solanum tuberosum (potato)
SOURCE         Solanum tuberosum
ORGANISM       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
               asterids; lamids; Solanales; Solanaceae; Solanum.
```

REFERENCE

```
AUTHORS        Provart,N.J., Ma,H., Willmitzer,L. and Mueller-Roeber,B.
TITLE          Cloning of 2 subunits of a G-protein from a potato guard cell cDNA
               bank
JOURNAL         Unpublished
REFERENCE      2 (bases 1 to 1524)
AUTHORS        Provart,N.J.
TITLE          Direct Submission
JOURNAL        Submitted (07-JUN-1995) N.J. Provart, Inst. fuer Genbiologische
               Forschung GmbH, Ihnestrasse 63, 14195 Berlin, FRG
```

FEATURES

source

```
1. .1524
   /organism="Solanum tuberosum"
   /mol_type="mRNA"
   /cultivar="Desiree"
   /db_xref="taxon:4113"
   /cell_type="guard-cell enriched"
   /tissue_type="epidermal fragment of leaf"
```

gene
CDS

```
/clone.lib="lambda ZapII"
/dev_stage="8 weeks"
1. .1524
/gene="gb1"
49. .1182
/gene="gb1"
/note="beta 1 subunit of heterotrimeric G-protein"
/codon_start=1
/product="GB1 protein"
/protein_id="CAA61106.1"
/db_xref="GI:1771734"
/db_xref="GOA:P93563"
/db_xref="InterPro:IPR001632"
/db_xref="InterPro:IPR001680"
/db_xref="InterPro:IPR011046"
/db_xref="UniProt/Swiss-Prot:P93563"
/translation="MSVAELKHERMAATQTVDNLRKLRKQRLQLDQDVSQYAKRQG
IKLPCAWMTCAFPSGQSVACGLDSACISIFNLNSPIDKDGIIHPVSMLSGHKGVYS
SCQYVDETHLITSSGQTCVLMDITTTGLRTSVFGGFQSGHTADVLVSISTSSNPK
LFVSGSDITAKLMDIRVASRAQRTTFHGHESDVTNKFPPDGNRRGTSGDSDGSLRFD
IRTGHLQVYNQPHGDDGDIPIHTSNMFSISGRLLFVGSNGSDCYVNDTLLAKVLNLG
SVQNSHEGRISICLGLSADGSALCTGSDNTNLKIWAFGGHRSVV"
```

ORIGIN

```
Query Match      39.3%; Score 640.8; DB 15; Length 1524;
Best Local Similarity 73.9%; Pred. No. 4.6e-172;
Matches 842; Conservative 0; Mismatches 292; Indels 6; Gaps 2;

Qy      205 GGAATGTCTGCTCCGAGCTCAAGAACGCCACGCGTCGCTACGAGACCGTTAATAAC 264
Db      46  GAAATGTCAATTGCGGAGCTGAAAGCGGCACATGGCCGCTACACAGACTGTAAATGAT 105
Qy      265 CTCCTGACCACTTTAGACAGACGCCCTCCAGCTCCTCGATACCGATCGGTGGAGGTAT 324
Db      106 CTCCTGAAAAACTTAAAGCAGAAGCGTCTCCAATTAATCTCGACACAGATGTTCTGGGTAT 165
Qy      325 TCAGCGGCGCAAGGAGTACTCGGTGAGCTTCGGAGCAACCGATCTGTTGTTGTCGT 384
Db      166 GCAAGAGCAAGGTAAAGTCCGGTAACTCCGGTAACTCCGCCCAACAGATCTAGTTGTTGATG 225
Qy      385 ACTCTTCAGGGACACACCGGAAAGTTTATTCAATTAGATTGGACACCGGAGAGAACCGG 444
Db      226 ATCTTGCAAGGACACACTGGAAGGTCTATTCTACGGAATGGAATCTCTGAAAAAATCGT 285
Qy      445 ATTGTGAGTGCATCTCAAGATCGGAGATTAATCGTGTGAATGCTCTAAACGAGTCAGAA 504
Db      286 ATAGTCAGTGCATCCCAAGATGGTAGATTAAATAGTGTGAATGCTCTCAAGCCAGAAA 345
Qy      505 ACTCATGCTATTAAACTCCCTTGTGCATGGGTTATGACATGCTCTTCTCTCCAAATGTT 564
Db      346 ACCCATGCAATTAAGCTTCCATGTGCTGGGTATGACACCTGTGCTCTCTCTAGTGA 405
Qy      565 CAGTGGTTCGCTGGTGGTATTTAGACAGTGTATGTTCTATCTTTAGCCTTAGCTCAACG 624
Db      406 CAGTCTGTTGCTTGTGGCGGCTTTGACAGTGCCTGCTCTATCTTCAACTTAATTCACCA 465
Qy      625 GGGGCAAGATGGAACTGTACCGGTTTCAAGAATGCTCACTGGTCACAGGGGATATGTT 684
Db      466 ATCGATAAGATGGGATCCATCCAGTATCGAAGATGCTTAGTGGGCATAAGGGGTATGTG 525
Qy      685 TCGTGTCTGATATGTCCCAATAGGATGCCACCTTATCAACAGTTTCAGTGTATCA 744
Db      526 TCTTCGTGTGATATGTTCCGGATGAGGATACCTACCTAATAACTAGTTCTCTGTTGATCAA 585
Qy      745 ACTTGATCTTATGGGATGTAACACTACTGCTCTCAAACTTCTGTTTTTGGCGGTGAATTT 804
Db      586 ACATGTGTACTTTGGGATATACTACTGGCCCTAAGAACTTCTGTGTTTGGAGGTGAGTTT 645
Qy      805 CAGTCTGGACATACCTGCTGTACTTAAGCGTCTCAATCAGTGGATCAAAACCAACTGG 864
Db      646 CAATCTGGGCACACTGTCAGATGTATTAAGTGTCTCAATAGTTTCACTTAACCCCAACTG 705
```

QY	865	TTTATATCTGTTTCATCGATTCCACAGCAGCGTTGTGGGACACTCGTGCTGCAAGCCGA	924
Db	706	TTTGTGTCGGGTCTGTGACACAACACTGCTGCAGTGTGGGACACCCCGAGTTGCTAGTCGA	765
QY	925	GCAGTGCCTACTCTTCATCGGTTCACGAGGAGATGTTAATACGTCACGTTCTTTCCGGAT	984
Db	766	GCTCAACGAACTTTCATGGAACAGAGAGTGAATTAATCTGTAAAGTTCTTCCTGAC	825
QY	985	GGGTATAGATTGGGATGGGATCAGACGATGAAACATCAGGCTGTATGACATAGGACT	1044
Db	826	GGTAAATAGATTGGAACTGGTTTCAGATGATGGAAGCTCAGATTATTGACATTAGGACT	885
QY	1045	GGTCAACCACTCCAGGT---CTATCAGGCACATGCTGATGTGAGAACGGACCTGTCAAC	1101
Db	886	GGACACCAAGCTGCAAGTATACAACCAACCGCATGGTGACGCTCCCTCATGTGACT	945
QY	1102	TCCATTGCAATCTCTGTCAGGAGACTTCTTTTCGCTGGCTATGCGGAGCAACAACACT	1161
Db	946	TCCATGGCAATTTCTATCTCAGGCGCTCTTCTCTTTGCGGGTACTC---TAATGGTGAT	1002
QY	1162	TGCTACGTTTGGGATACCTCTTGGGAGAGTTGTTATGGATTGCGGATTAACAGCAGGAT	1221
Db	1003	TGTTACGTGTGGGACACCTATTAGCAAGGTGGTCTTAAACCTTAGGATCAGTTCAAAAC	1062
QY	1222	TCACACAGGAATAGATTAAGCTGTTTGGGTTGTTCAGCAGATGGAAGTGCAATTGTGTACA	1281
Db	1063	TCTCATGAAGGCGGAATAAGTTGTCCTGGGACTGTGAGCTGACGGAAGTGCCCTTATGTACA	1122
QY	1282	GGAGTTGGGATTCAAACTTAAAGATATGGGCGTTTGGGAGCACACAGGAGTGATTGA	1341
Db	1123	GGAGTTGGGATACAAACCTGNAGATTGGGCTTTTGGAGGACACAGAAGTGTGCTCTGA	1182
RESULT 15			
AF145976			
LOCUS			
Pisum sativum G protein beta subunit mRNA, complete cds.			
AF145976			
AF145976.1 GI:4929351			
KEYWORDS			
Pisum			
SOURCE			
Pisum sativum (pea)			
ORGANISM			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.			
REFERENCE			
1 (bases 1 to 1611)			
Wu, Y. and Tuteja, N.			
Isolation and characterization of G protein from pea			
JOURNAL			
Unpublished			
REFERENCE			
2 (bases 1 to 1611)			
Wu, Y. and Tuteja, N.			
Direct Submission			
Submitted (26-APR-1999) Plant Molecular Biology, ICGEB-New Delhi,			
JOURNAL			
Aruna Asaf Ali Marg, New Delhi 110067, India			
FEATURES			
source			
1..1611			
/organism="Pisum sativum"			
/mol_type="mRNA"			
/db_xref="taxon:3888"			
144..1277			
/function="involved as a modulator or transducer in various transmembrane signaling systems"			
/note="contains 7 WD repetitive region"			
/codon_start=1			
/product="G protein beta subunit"			
/protein_id="AAD33959.1"			
/db_xref="GI:4929352"			
/translation="MSVADVKEHIAATETVNNLRSLRDLISLLDITDIAGYARSQG RAPVFGPTDLICRQLAGTKVYSLDWTSEKRIVSASQDGRILVNNALTSOKTHA IKLPCAWMTCAFSPTGSGVACGLDSVCSIFNLNSPLDRGNLNVSMLSGHKGYVS SCQVYVGRDHLITGSGDQTCVLMDDITTLGLRATSVLGRFSQGHADVLISINGSNSK LFSVSGSCDARLMDTRVASRAVTFPHGHEGVNSKVFPPGNRFGTSGEDGTCRLFPD			
CDS			
1091 GACCTGTCACTCCATTGCATTTCTCTGTGTCAGGAGACTTCTTTTCGCTGGCTATCGGA			
1030 CACATGTGACGCTCCATTGCATTTTCCATATCCCGAAGACTTCTTATTGCTGGCTAT--A			
1151 GCAACAACACTTGTCTAGCTTTGGGATACCTCTTGGGAGAGGTTGTATTGGATTGGAT			
1210			
Query Match			
Best Local Similarity 39.3%; Score 640.6; DB 15; Length 1611;			
Matches 846; Conservative 0; Mismatches 239; Indels 6; Gaps 2;			
ORIGIN			
QY	194	AGTCTGATCCCGGAATGTCTGTCTCCGAGCTCAAAGAACCCACGCGTCCGTACCGGAGA	253
Db	130	ATTCACTTGACAAATGTCGGTTCGGGACGTCAAAGAACGTCACATAGCAGCGAGAA	189
QY	254	CCGTTAATAACCTCCGTGACACAGCTTAGACAGAGACGCTCCAGCTCTCTCATACCGATG	313
Db	190	CGGTTAAACAATCTCAGAGAACGATTGAGCAGAGACCGGCTTTCTTTGCTGATACAGATA	249
QY	314	TGGCGAGGATTTCAGCGCGCAAGGACGTACTCGGGTGAGCTTCGGAGCAACGATCTGG	373
Db	250	TTGCTGGATATGCTAGGTCTCAAGGTAGAGTCTCTGTATCTTTTGGTCCCACTGATATTC	309
QY	374	TTTGTGTGCTACTCTTCAGGGAACACACCGGAAAGGTTTATTCATTAGATTGACACCGG	433
Db	310	TTTGTGTAGAACGCTCCAAGGTCATACCGGAAAGGTGATTTCATTGGATTGGACTTCAG	369
QY	434	AGAGAAACCGGATTTGTCAGTGCACTCAAGATGGGAGATTATTCGTGTGGAATGCTCTAA	493
Db	370	AAAAGAAATAGGATTTGTTAGTGCAATCCCAAGATGGAAGATTAAATAGTGTGGAATCTCTAA	429
QY	494	CGAGTCAGAAACCTCATGCTATTAAACCTCCCTCTGTGATGGGTTATGACATGCTCTTCT	553
Db	430	CAAGCCAGAAACCTCATGCTATAAAGCTTCTTGTGCAATGGGCTGATGACGCTGCTTCT	489
QY	554	CTCCAAATGGTTCAGTCCGTTGCGTGTGGATTAGACAGGTATGTATCTTCTTTAGCC	613
Db	490	CACCAACTGGTCAATCTGTTGCTTGTGGGSCCTTGACAGTGTGTTGCTCTATTTCATC	549
QY	614	TTAGCTCAACGGCGGACAGAGATGGAATGTACCGGTTTCAAGAAATGCTCACTGCTGACA	673
Db	550	TTAAATCTCCACTCGATAGGATGGGAATCTAAATGTTTTCACGGATGCTTAGTGGACATA	609
QY	674	GGGATATGTTTTCGTCGTCTCAGTATGTCCTCAAAATGAGGATGCCACCTTATCACCAGT	733
Db	610	AAGTTATGTTTCACTTGTGCAATGTTTCAAGTGAAGACACTCACTTATCACTGTT	669
QY	734	CAGGTGATCAAACTGTATCTTATGGGATGTAACTACTGCTCTCAAAACCTTCTGTTTTG	793
Db	670	CTGAGATCAGACATGTGTTTATGCGATATTACTACTGGCTTTAGAACAATCTGCTTTT	729
QY	794	CGGTGAATTTTCAGTCTGGACATACCTGCTGATGTACTAAGGCTCTCAATCAGTGCATCA	853
Db	730	TAGCGAGTTTTCAGTCTGGACATACCTGAGATGTACTTAGCATTTCCATTAAATGGATCCA	789
QY	854	ACCCAAACTGGTTTATATCTGTTTCATGCGATCCACAGCAGGTTGTGGGACACTCGTG	913
Db	790	ACTCCAAATGTTTGTATCTGGTTCTTGGATGCGATGCCAGATTTGGGACACTCGTG	849
QY	914	CTGCAAGCCGAGCAGTGGTACCTTTTCATGCTCAGAGGAGATGTTTAATACCGTCAAGT	973
Db	850	TGGCAAGTCGAGCAGTGGGACATTTTCAGCGCCACGAGGGAGATGTTTAAATTTCTGCAAT	909
QY	974	TCTTTCCGATGGGTATAGATTGGGATCGATCGATGGAACATGCGAGGCTGTATG	1033
Db	910	TCTTTCCGATGGAAATAGATTGGAACTGGCTCAGAGGATGGAATTTGCAATTTATTTG	969
QY	1034	ACATAAGGACTGGTCCCAACTCCAGGTCT---ATCAGCCACATGCTGATGGTGCAGAACG	1090
Db	970	ACATTAGACCGGACACCAACTTCAAGTATATTAATCAGAACCAACAGAACCAAGATGG	1029
QY	1091	GACCTGTCACTCCATTGCATTTCTCTGTGTCAGGAGACTTCTTTTCGCTGGCTATCGGA	1150
Db	1030	CACATGTGACGCTCCATTGCATTTTCCATATCCCGAAGACTTCTTATTGCTGGCTAT--A	1086
QY	1151	GCAACAACACTTGTCTAGCTTTGGGATACCTCTTGGGAGAGGTTGTATTGGATTGGAT	1210

```

Db      1087 CAAATGGTGATTGCTATGTTGGGATACCTTTATTGGCTAAGGTGGTCTTGAATCTAGGAT 1146
Qy      1211 TACAGCAGGATTTCACACAGGMAATAGAATAAGCTGTTTGGGTTTGTACGCAGATGGAAGTG 1270
Db      1147 CTCCTTCAAAACTCTCATGAGGSCAGGATCACCTGTTTGGGTATGTCGTGATGGAAGCG 1206
Qy      1271 CATTGTGTACAGGAAGTTGGGATTCAAATCTAAAGATATGGCCGTTTGGAGGACACAGGA 1330
Db      1207 CTTTATGTACAGGAAGTTGGGACACAAATTTAAAGATATGGGCAATTTGGAGGGCATAGGA 1266
Qy      1331 GAGTGATTTGA 1341
Db      1267 AGGTGATTTGA 1277

```

Search completed: December 28, 2005, 17:47:36
 Job time : 8039 secs

THIS PAGE BLANK (user)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: December 28, 2005, 08:46:01 ; Search time 943 Seconds
(without alignment)
11513.027 Million cell updates/sec

Title: US-10-602-898A-1
Perfect score: 1629
Sequence: 1 cctgacgtacgctgttg.....cagaagataaaacgctacta 1629

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*
14: Geneseqn2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1629	100.0	1629	12	Adh78721 Arabidops
2	1439.4	88.4	1465	3	Aac45984 Arabidops
3	1434.6	88.1	1467	3	Aac37381 Arabidops
4	1132.4	69.5	1134	12	Adn72944 Thale cre
5	728.6	44.7	802	6	Abn99039 Arabidops
6	661.6	40.6	1526	12	Adh78737 Pea AGB1
7	656.4	40.3	1430	12	Adh78735 Tobacco A
8	654.8	40.2	1600	12	Adh78729 Tobacco A
9	651	40.0	1434	12	Adh78733 Tobacco A
10	647.4	39.7	1676	14	Aeb26929 Pinus rad
11	642.4	39.4	1540	12	Adh78725 Potato AG
12	641.2	39.4	1560	12	Adh78731 Potato AG
13	640.8	39.3	1524	12	Adh78727 Potato AG
14	640.6	39.3	1611	12	Adh78739 Pea AGB1
15	637.6	39.1	693	3	Aac44600 Arabidops
16	576.4	35.4	1664	12	Adm47702 Polynucle
17	558.8	34.3	1480	12	Adm47702 Polynucle
18	558.8	34.3	1671	12	Adh78745 Maize AGB
19	558.2	34.3	1924	14	Aeb27029 Pinus rad

20	551.8	33.9	1470	12	ADH78741	Wild oat
21	545	33.5	756	3	AAC43814	Arabidops
22	529.8	32.5	1123	13	ADX27418	Plant ful
23	433	26.6	984	12	ADJ39635	Plant cdn
24	322.4	19.8	594	12	ADJ44713	Plant cdn
25	222.2	13.6	786	3	AAC51866	Arabidops
26	142	8.7	2208	12	ADG73724	Aspergill
27	138	8.5	1457	4	AAI59460	Human pol
28	138	8.5	2567	4	AHH14842	Human cdn
29	138	8.5	2567	10	ADD18943	Human dis
30	138	8.5	2567	13	ADR24374	Breast ca
31	138	8.5	2923	10	ADD12690	Human cdn
32	136.6	8.4	1023	9	ADA02731	Mouse Gnb
33	136.6	8.4	1023	10	ADB72469	Mouse Gnb
34	136.6	8.4	1023	10	ADE95979	Mouse Gnb
35	136.6	8.4	1410	6	AB199519	Mouse isc
36	136.6	8.4	1410	9	ADA02730	Mouse Gnb
37	136.6	8.4	1410	10	ADB72468	Mouse Gnb
38	136.6	8.4	1410	10	ADE95978	Mouse DNA
39	136.6	8.4	1410	10	ABV75223	Murine GN
40	135	8.3	2837	10	ADB58549	Toxicity-
41	135	8.3	2837	10	ADB53156	Primary r
42	135	8.3	2837	10	ABT42151	Toxicity
43	135	8.3	2837	13	ADV41343	Rat card1
44	131	8.0	669	12	ADJ43948	Plant cdn
45	130	8.0	1520	3	AAA74590	Rat Taste

ALIGNMENTS

RESULT 1

ID ADH78721 standard; DNA; 1629 BP.

AC ADH78721;

DT 22-APR-2004 (first entry)

DE Arabidopsis thaliana AGB1 gene, SEQ ID 1.

KW Plant; agronomic trait; flowering; fruit yield; seed yield; root biomass;
KW seed size; seed shape; stem branch number; plant size; gene; ds;
KW GPA.

OS Arabidopsis thaliana.

PN WO2004003146-A2.

PD 08-JAN-2004.

PF 24-JUN-2003; 2003WO-US020001.

PR 28-JUN-2002; 2002US-0392730P.

PR 05-FEB-2003; 2003US-0445208P.

PA (PARA-) PARADIGM GENETICS INC.
(UYNC-) UNIV NORTH CAROLINA.

PI Boyes D, Davis K, Jones A, Ullah H, Chen J, Mulpuri R;
PI Chatterjee A, Ward MP;

DR WPI; 2004-083031/08.

DR P-PSDB; ADH78722.

PT Altering a plant agronomic trait, useful for improving plant agronomic
traits comprising regenerating plants having stably integrated expression
cassette, where the regenerated plant has an altered agronomic trait.

PS Claim 4; SEQ ID NO 1; 180pp; English.

XX The present invention relates to a method for altering a plant agronomic
trait selected from time to flowering, duration of flowering, fruit

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 48488.
XX DE
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX XX
OS Arabidopsis thaliana.
XX XX
PN EP1033405-A2.
XX XX
PD 06-SEP-2000.
XX XX
PF 25-FEB-2000; 2000EP-00301439.
XX XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126364P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-012834P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144353P.
PR 21-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 26-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.

PR	27-AUG-1999;	99US-0151080P.	
PR	30-AUG-1999;	99US-0151303P.	
PR	31-AUG-1999;	99US-0151438P.	
PR	01-SEP-1999;	99US-0151930P.	
PR	07-SEP-1999;	99US-0152363P.	
PR	10-SEP-1999;	99US-0153070P.	
PR	13-SEP-1999;	99US-0153758P.	
PR	15-SEP-1999;	99US-0154018P.	
PR	16-SEP-1999;	99US-0154039P.	
PR	20-SEP-1999;	99US-0154779P.	
PR	22-SEP-1999;	99US-0155139P.	
PR	23-SEP-1999;	99US-0155486P.	
PR	24-SEP-1999;	99US-0155659P.	
PR	28-SEP-1999;	99US-0156458P.	
PR	29-SEP-1999;	99US-0156596P.	
PR	04-OCT-1999;	99US-0157117P.	
PR	05-OCT-1999;	99US-0157553P.	
PR	06-OCT-1999;	99US-0157865P.	
PR	07-OCT-1999;	99US-0158029P.	
PR	08-OCT-1999;	99US-0158232P.	
PR	12-OCT-1999;	99US-0158369P.	
PR	13-OCT-1999;	99US-0159293P.	
PR	13-OCT-1999;	99US-0159294P.	
PR	13-OCT-1999;	99US-0159295P.	
PR	14-OCT-1999;	99US-0159329P.	
PR	14-OCT-1999;	99US-0159330P.	
PR	14-OCT-1999;	99US-0159331P.	
PR	14-OCT-1999;	99US-0159637P.	
PR	14-OCT-1999;	99US-0159638P.	
PR	18-OCT-1999;	99US-0159584P.	
PR	21-OCT-1999;	99US-0160741P.	
PR	21-OCT-1999;	99US-0160767P.	
PR	21-OCT-1999;	99US-0160768P.	
PR	21-OCT-1999;	99US-0160770P.	
PR	21-OCT-1999;	99US-0160814P.	
PR	21-OCT-1999;	99US-0160815P.	
PR	22-OCT-1999;	99US-0160980P.	
PR	22-OCT-1999;	99US-0160981P.	
PR	22-OCT-1999;	99US-0160989P.	
PR	25-OCT-1999;	99US-0161404P.	
PR	25-OCT-1999;	99US-0161405P.	
PR	25-OCT-1999;	99US-0161408P.	
PR	26-OCT-1999;	99US-0161559P.	
PR	26-OCT-1999;	99US-0161360P.	
PR	26-OCT-1999;	99US-0161361P.	
PR	28-OCT-1999;	99US-0161920P.	
PR	28-OCT-1999;	99US-0161992P.	
PR	28-OCT-1999;	99US-0161993P.	
PR	29-OCT-1999;	99US-0162142P.	
Query Match 88.4%; Score 1439.4; DB 3; Length 1465;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1451; Conservative 0; Mismatches 1; Indels 1; Gaps 1;			
Qy	13	CGTGTTCGTCTTGACATGATTCCTCTCAAGCTTTTAAATCTCTCTCTTTTCCCA	72
Db	14	CGTGTTCGTCTTGACATGATTCCTCTCAAGCTTTTAAATCTCTCTCTTTTCCCA	73
Qy	73	CGTAATCCCCCAATCCATCTTTCTAGGGTTGATCTCCCTCTCTCAATCATGAACCT	132
Db	74	CGTAATCCCCCAATCCATCTTTCTAGGGTTGATCTCCCTCTCTCAATCATGAACCT	133
Qy	133	TCCTCTCTCTAGACCCCAAAAGTTTCCCTCTTTATTTGATCGGCGACGGAAGGCT	192
Db	134	TCCTCTCTCTAGACCCCAAAAGTTTCCCTCTTTATTTGATCGGCGACGGAAGGCT	193
Qy	193	AAGCTGATCCCGAATGCTCTCCGAGCTCAAAGAACCGCGCGCTGCTACGGAG	252
Db	194	AAGCTGATCCCGAATGCTCTCTCCGAGCTCAAAGAACCGCGCGCTGCTACGGAG	253
Qy	253	ACCGTTAATAACCTCCGATGACAGCTTAGACAGACAGCCCTCCAGCTCCTCGATACCGAT	312
Db	254	ACCGTTAATAACCTCCGATGACAGCTTAGACAGACAGCCCTCCAGCTCCTCGATACCGAT	313

Qy	313	GTGGCGAGGTATTACAGCGCGCAAGGACGTACTCGGTGAGCTTTCGGAGCAACGGATCTG	372
Db	314	GTGGCGAGGTATTACAGCGCGCAAGGACGTACTCGGTGAGCTTTCGGAGCAACGGATCTG	373
Qy	373	GTTTGTTCGTACTCTTCAGGGACACACCGAAAGTTTATTATTAGATTGGACACCG	432
Db	374	GTTTGTTCGTACTCTTCAGGGACACACCGAAAGTTTATTATTAGATTGGACACCG	433
Qy	433	GAGAGGAACCGGATTTGTCAGTGCAATCTCAAGATCGGAGATTAATCGTGTGGAATGCTCTA	492
Db	434	GAGAGGAACCGGATTTGTCAGTGCAATCTCAAGATCGGAGATTAATCGTGTGGAATGCTCTA	493
Qy	493	ACGAGTCAAGAAACATCATGCTATTAAACTCCCTTTGTCATGGGTTATGACATGTGCTTTC	552
Db	494	ACGAGTCAAGAAACATCATGCTATTAAACTCCCTTTGTCATGGGTTATGACATGTGCTTTC	553
Qy	553	TCTCCAAATGGTCAAGTGGTTGCGTGTGGTGGATAGACAGTGTATGTTCTATTTCAGC	612
Db	554	TCTCCAAATGGTCAAGTGGTTGCGTGTGGTGGATAGACAGTGTATGTTCTATTCTTTAGC	613
Qy	613	CTTAGCTCAACCGCGGACCAAGATGGAACCTGTACCGGTTTCAAGAAATGCTCACTGGTCAC	672
Db	614	CTTAGCTCAACCGCGGACCAAGATGGAACCTGTACCGGTTTCAAGAAATGCTCACTGGTCAC	673
Qy	673	AGGGGATATGTTTCGTGCTGTCAGTATGTCCCAAATGAGGATGCCACCTTATCACCAGT	732
Db	674	AGGGGATATGTTTCGTGCTGTCAGTATGTCCCAAATGAGGATGCCACCTTATCACCAGT	733
Qy	733	TCAGGTGATCAAACTTGTTGATCTTATGGGATGTAATCTACTGCTCTCAAACTTCTGTTTTT	792
Db	734	TCAGGTGATCAAACTTGTTGATCTTATGGGATGTAATCTACTGCTCTCAAACTTCTGTTTTT	793
Qy	793	GGCGGTGAATTTCAAGTCTGGACATCTGCTACTTAAGCGCTCTCAATCAGTGGATCA	852
Db	794	GGCGGTGAATTTCAAGTCTGGACATCTGCTACTTAAGCGCTCTCAATCAGTGGATCA	853
Qy	853	AACCCAAACTGGTTTATATCTGTTTCATGCGATTCCACAGCACCGTTGTGGGACACTCGT	912
Db	854	AACCCAAACTGGTTTATATCTGTTTCATGCGATTCCACAGCACCGTTGTGGGACACTCGT	913
Qy	913	GCTGCAAGCCGAGCAGTGGTACCTTTCATGCTGTCACGAGGAGATGTTAATACGGTCAAG	972
Db	914	GCTGCAAGCCGAGCAGTGGTACCTTTCATGCTGTCACGAGGAGATGTTAATACGGTCAAG	973
Qy	973	TTCTTTCCGGATGGTATAGATTTGGGACTGCGATCAGACGATGGAACATGCAAGGCTGAT	1032
Db	974	TTCTTTCCGGATGGTATAGATTTGGGACTGCGATCAGACGATGGAACATGCAAGGCTGAT	1033
Qy	1033	GACATAAGGACTGGTCAACCACTCCAGGTCTATCAGCCACATGGTGTGAGAACCGGA	1092
Db	1034	GACATAAGGACTGGTCAACCACTCCAGGTCTATCAGCCACATGGTGTGAGAACCGGA	1093
Qy	1093	CCTGTACCTCCATTCGATTCCTGTGTACGGGAGACCTTCTTTTCGTGGCTATGCGAGC	1152
Db	1094	CCTGTACCTCCATTCGATTCCTGTGTACGGGAGACCTTCTTTTCGTGGCTATGCGAGC	1153
Qy	1153	AACAACACTTGTCTACGTTTGGGATACCTCTTTGGGAGAGGTTGATTTGGATTTGGGATTA	1212
Db	1154	AACAACACTTGTCTACGTTTGGGATACCTCTTTGGGAGAGGTTGATTTGGATTTGGGATTA	1213
Qy	1213	CAGCAGGATTCACACAGGAATAGAAATGAGCTGTTTGGGTTTGTGAGGATGGAAGTGC	1272
Db	1214	CAGCAGGATTCACACAGGAATAGAAATGAGCTGTTTGGGTTTGTGAGGATGGAAGTGC	1273
Qy	1273	TTGTGTACAGGAAGTTGGGATTCAAATCTAAAGATATGGGCGTTTGGAGGACAGAGGA	1332
Db	1274	TTGTGTACAGGAAGTTGGGATTCAAATCTAAAGATATGGGCGTTTGGAGGACAGAGGA	1333
Qy	1333	GTGATTTGAAGAAGATTTAAACGAAAGTAGGAGTCACTCTCCAGTTGTTTGGTTAATA	1392
Db	1334	GTGATTTGAAGAAGATTTAAACGAAAGTAGGAGTCACTCTCCAGTTGTTTGGTTAATA	1392

```
QY 1393 TTCTGTAGTCGGAGTAAAGTTTCGGTTTGTGGAAGGTGTTTGGTTTGAATAGTGAGT 1452
DB 1393 TTCTGTAGTCGGAGTAAAGTTTCGGTTTGTGGAAGGTGTTTGGTTTGAATAGTGAGT 1452
QY 1453 GGTAGAGAAATT 1465
DB 1453 GGTAGAGAAATT 1465

RESULT 3
AAC37381
ID AAC37381 standard; DNA; 1467 BP.
XX
AC AAC37381;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17173.
XX
DE Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 21-MAY-1999; 99US-0135124P.
PR 24-MAY-1999; 99US-0135353P.
PR 25-MAY-1999; 99US-0135629P.
PR 27-MAY-1999; 99US-0136021P.
PR 28-MAY-1999; 99US-0136392P.
PR 01-JUN-1999; 99US-0136782P.
PR 03-JUN-1999; 99US-0137222P.
PR 04-JUN-1999; 99US-0137528P.
PR 07-JUN-1999; 99US-0137502P.
PR 08-JUN-1999; 99US-0137724P.
PR 10-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140151P.
PR 23-JUN-1999; 99US-0140254P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0146389P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
```

[illegible]

Db 841 CACCAACTCCAGGCTCTATCAGCCACATGTTGATGTGAGAACGGACCTGTCCACCTCCATT 900
QY 1108 GCATTCCTCTGTCTCAGGAGACTCTCTTTTCGCTGCTATGCCAGCAACAACTTGTCTAC 1167
Db 901 GCATTCCTCTGTCTCAGGAGACTCTCTTTTCGCTGCTATGCCAGCAACAACTTGTCTAC 960
QY 1168 GTTGGGATACCTCTTGGGAGAGGTTGTATTTGGATTTGGGATTTACAGAGGATTCACAC 1227
Db 961 GTTGGGATACCTCTTGGGAGAGGTTGTATTTGGATTTGGGATTTACAGAGGATTCACAC 1020
QY 1228 AGGAATAGAAATAGCTGTTTGGGGTTGTCAGCAGATGGAAGTGCATTTGTACAGGAAGT 1287
Db 1021 AGGAATAGAAATAGCTGTTTGGGGTTGTCAGCAGATGGAAGTGCATTTGTACAGGAAGT 1080
QY 1288 TGGGATTCAAATCTAAAGATATGGGCGTTTGGAGGACACAGGAGAGTGAATTGA 1341
Db 1081 TGGGATTCAAATCTAAAGATATGGGCGTTTGGAGGACACAGGAGAGTGAATTGA 1134

RESULT 5
ID ABN99039
XX ABN99039 standard; DNA; 802 BP.
AC ABN99039;
XX
DT 01-AUG-2002 (first entry)
XX
DE Arabidopsis thaliana expressed polynucleotide SEQ ID NO 807.
XX
KW Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
KW disease; crop; thale cress; tolerance factor; insect; pathogen;
KW nutrition; ds.
XX
OS Arabidopsis thaliana.
XX
PN US2002023281-A1.
XX
PD 21-FEB-2002.
XX
PF 26-JAN-2001; 2001US-00770445.
XX
PR 27-JAN-2000; 2000US-0178472P.
XX

(GORL/) GORLACH J.
PA (ANY/) AN Y.
PA (HAMI/) HAMILTON C M.
PA (PRIC/) PRICE J L.
PA (RAIN/) RAINES T M.
PA (YUY/) YU Y.
PA (RAME/) RAMEAKA J G.
PA (PAGE/) PAGE A.
PA (MATH/) MATHAW A V.
PA (LEDF/) LEDFORD B L.
PA (WOES/) WOESSNER J P.
PA (HAAS/) HAAS W D.
PA (GARC/) GARCIA C A.
PA (KRIC/) KRICKER M.
PA (SLAT/) SLATER T.
PA (DAVI/) DAVIS K R.
PA (ALLE/) ALLEN K.
PA (HOFF/) HOFFMAN N.
PA (HURB/) HURBAN P.
XX
PI Gorlach J, An Y, Hamilton CM, Price JL, Raines TM, Yu Y;
PI Rameaka JG, Page A, Mathew AV, Ledford BL, Woessner JP, Haas WD;
PI Garcia CA, Kricker M, Slater T, Davis KR, Allen K, Hoffman N;
PI Hurban P;
XX
DR WPI; 2002-403163/43.
XX
XX New Arabidopsis thaliana nucleic acid for identifying homologous genes,
PT producing compositions that modulate the expression or function of its

PT encoded protein, and mapping functional regions of protein.
XX
PS Claim 1; SEQ ID NO 807; 49pp + Sequence Listing; English.
XX
CC The invention relates to an Arabidopsis thaliana nucleic acid (I)
CC comprising a sequence capable of hybridising under stringent conditions
CC to a sequence selected from any one of 999 sequences (ABN98233-ABN99231),
CC given in the specification or its fragment. A polypeptide (II) encoded by
CC (I), a transgenic plant (III) comprising an exogenous nucleic acid or a
CC genetically modified cell (IV) comprising an exogenous nucleic acid, is
CC useful for screening a candidate agent for its biological effect. (I) is
CC useful in identifying homologous or related genes, in producing
CC compositions that modulate the expression or function of its encoded
CC protein, mapping functional regions of the protein and in studying
CC associated physiological pathways. (I) is also useful for the genetic
CC manipulation of cells, particularly plant cells. (I) is also useful in
CC screening assays of various plant strains to determine the strains that
CC are best capable of withstanding a particular disease or environmental
CC stress. (II) and (III) are useful for screening of biologically active
CC agents, e.g. fungicides, insecticides, etc., for elucidating biochemical
CC pathways. The screened agents are useful in improved methods of treating
CC crops to prevent or treat disease. (II) are also useful in screening
CC programs to identify agents that mimic or enhance the action of tolerance
CC factors. Such agents are useful in improved methods of treating crops to
CC enhance their tolerance to environmental stress. (I) is also useful for
CC enhancing or inhibiting production of a biosynthetic product in a plant.
CC (III) is useful for identifying other mediators that may induce
CC expression of proteins of interest, for establishing the extent to which
CC any specific insect and/or pathogen is responsible for damage to a
CC particular plant, for identifying other mediators that enhance or induce
CC tolerance to environmental stress, for identifying factors involved in
CC biosynthetic pathways of nutritional, commercial, or medicinal value and
CC for identifying productions of nutritional, commercial or medicinal
CC value. (IV) is useful in the study of genetic function and regulation,
CC for alteration of the cellular metabolism and for screening compounds
CC that may affect the biological function of the gene or gene products.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from USPTO
XX at seqdata.uspto.gov/sequence.html?docID=99909770445
XX
SQ Sequence 802 BP; 230 A; 133 C; 202 G; 227 T; 0 U; 10 Other;

Query Match 44.7%; Score 728.6; DB 6; Length 802;
Best Local Similarity 95.5%; Pred. No. 2.2e-215;
Matches 765; Conservative 0; Mismatches 34; Indels 2; Gaps 2;
QY 821 CTGATGTACTAAGCGTCTCAATCAGTGGATCAAAACCCAACTGGTTATATCTGTTTCAT 880
Db 1 CTGATGTACTAAGCGTCTC-ATCAGTGGATCAAAACCCAACTGGTTATATCTGTTTCAT 59
QY 881 CGGATTCACAGCAGCGTTGTGGACACTCGTCTGCAAGCCGAGCAGTGCCTACCTTTC 940
Db 60 CGGATTCACAGCAGCGTTGTGGACACTCGTCTGCAAGCCGAGCAGTGCCTACCTTTC 119
QY 941 ATGCTCAGGAGGAGATGTTAATACGTTCTTTCCGGATGGGTATAGATTGGGA 1000
Db 120 ATGCTCAGGAGGAGATGTTAATACGTTCTTTCCGGATGGGTATAGATTGGGA 179
QY 1001 CTGGATCAGACGATGGAACATGCAAGGCTGTATGACATAAGGACTGGTCACCACTCCAGG 1060
Db 180 CTGGATCAGACGATGGAACATGCAAGGCTGTATGACATAAGGACTGGTCACCACTCCAGG 239
QY 1061 TCTATCAGCCACATGGTGATGGTGAGAACCGAACCTGTCCACTCCATTCCTCTGTGT 1120
Db 240 TCTATCAGCCACATGGTGATGGTGAGAACCGAACCTGTCCACTCCATTCCTCTGTGT 299
QY 1121 CAGGAGACTTCTTTTTCGCTGGCTATGCGAGCAACACACTTGTACGTTTGGGATACCC 1180
Db 300 CAGGAGACTTCTTTTTCGCTGGCTATGCGAGCAACACACTTGTACGTTTGGGATACCTN 359
QY 1181 TCTTGGGAGAGTGTATTGGATTGGATTACAGCAGGAGATTACACAGGAAATAGATAA 1240
Db 360 TCTTGGGAGAGTGTATTGGATTGGATTACAGCAGGAGATTACACAGGAAATAGATAA 419

1241 GCTGTTGGGTTGTCAGCAGATGGAGTGCATTTGTGTACAGGAAGTTGGGATTCAAATC 1300
 420 GCTGTTGGGTTGTCAGCAGATGGAGTGCCTTTGTGTACAGGAAGTTGGGATTCAAATC 479
 1301 TAAAGATATGGGCGTTTGGAGGACACAGCAGAGTGAATTTGAAGAAGATTTACGAAAGT 1360
 480 TAAAGATATGGGCGTTTGGAGGACACAGCAGAGTGAATTTGAAGAAGATTTACG- AAGT 538
 1361 AGGAGTCAGCTCTCCAGTTTGGTTTAAATATATTTCTGTAGTCGGGAAGTAAGTTTCGGTT 1420
 539 AGGAGTCACNNNTCCAGTTTGGTTTAAATATANNNGTAGTCGGGAAGTAAGTTTCGGTT 598
 1421 TGTGGAAGTGTGTTGGTTTGAATAGTAGGAGTGTAGAGAAATTAACCTTCCCTTTTGG 1480
 599 TGTGGAAGTGTGTTGGTTTGAATAGTAGGAGTGTAGAGAAATTAACCTTCCCTTTTGG 658
 1481 TAGTGTGCTTTGATTTATTTATTTCTTCAATTTGGAACTAACTCCCTTCAACACGCTACTC 1540
 659 TAGTGTGCTTTGATTTATTTATTTCTTCAATTTGGAACTAACTCCCTTCAACACGCTACTC 718
 1541 AATGTGAATTTCTGTAAATCAATTTGTGTACCCAGCTCTTTTACTTACTATCTCTTCA 1600
 719 AATGTGAATTTCTGTAAATCAATTTGTGTACCCAGCTCTTTTACTTAAATTAATTAATTA 778
 1601 TATTGAACGAGAAGATAAAA 1621
 779 AAAAAAAAAAAAAAAAAAAAAA 799

RESULT 6

ADH78737
 ID ADH78737 standard; DNA; 1526 BP.
 AC ADH78737;
 DT 22-APR-2004 (first entry)
 XX Pea AGB1 gene ortholog, SEQ ID 17.
 DE Plant; agronomic trait; flowering; fruit yield; seed yield; root biomass;
 KW seed size; seed shape; stem branch number; plant size; AGB1; gene; ds;
 KW Pea; GPA.
 XX Pisum sativum.
 OS WO2004003146-A2.
 PN 08-JAN-2004.
 PD 24-JUN-2003; 2003WO-US020001.
 PP 28-JUN-2002; 2002US-0392730P.
 PR 05-FEB-2003; 2003US-0445208P.
 XX (PARA-) PARADIGM GENETICS INC.
 PA (UYNC-) UNIV NORTH CAROLINA.
 XX Boyes D, Davis K, Jones A, Ullah H, Chen J, Mulpuri R;
 PI Chatterjee A, Ward MP;
 PI WPI; 2004-083031/08.
 DR P-PSDB; ADH78738.
 XX Altering a plant agronomic trait, useful for improving plant agronomic
 PT traits comprising regenerating plants having stably integrated expression
 PT cassette, where the regenerated plant has an altered agronomic trait.
 XX Disclosure; SEQ ID NO 17; 180pp; English.
 XX The present invention relates to a method for altering a plant agronomic
 CC trait selected from time to flowering, duration of flowering, fruit
 CC yield, seed yield, root biomass, seed size, seed shape, number of stem

branches or size of plant. The method for altering a plant agronomic
 CC trait comprises: introducing into a plant cell an expression cassette
 CC comprising a nucleotide sequence operably linked to a promoter that is
 CC operable within the plant cell, where the nucleotide sequence is selected
 CC from: a nucleotide sequence antisense to a plant AGB1 or an AGB1 ortholog
 CC (ADH78721-ADH78722 and ADH78725-ADH78746); a nucleotide sequence
 CC comprising an inverted repeats of AGB1 or an AGB1 ortholog; a nucleotide
 CC sequence encoding a dsRNA comprising a first RNA complementary to at
 CC least 25 consecutive nucleotides of a plant AGB1 or an AGB1 ortholog and
 CC a second RNA substantially complementary to the first RNA; a nucleotide
 CC sequence that is AGB1 or an AGB1 ortholog; or a nucleotide sequence that
 CC is GPAL or a GPAL ortholog (ADH78723-ADH78724 and ADH78747-ADH78781); and
 CC regenerating a plant that has a stably integrated expression cassette
 CC from the plant cell, where the regenerated plant has an altered agronomic
 CC trait. The method is useful for improving plant agronomic traits
 CC including duration of flowering, fruit yield, root biomass, seed size,
 CC seed shape, number of stem branches or plant size. The transgenic plants
 CC obtained are useful as parents to produce progeny plants and plant lines.
 XX
 SQ Sequence 1526 BP; 416 A; 268 C; 352 G; 490 T; 0 U; 0 Other;
 Query Match 40.6%; Score 661.6; DB 12; Length 1526;
 Best Local Similarity 72.8%; Pred. No. 2.2e-194;
 Matches 882; Conservative 0; Mismatches 324; Indels 6; Gaps 2;
 QY 207 AATGTCGTCTCCGAGCTCAAAAGACGCCCGTCGTACGAGACCGTTAATAACCT 266
 DB 51 AATGTCGTCTCCGAGCTCAAAAGACCGTCGTACGAGACCGTTAATAACCT 110
 QY 267 CGTCACACAGCTTACAGACAGACGCTCCAGCTCTCGATACCGATGTGGCAGGTATTC 326
 DB 111 CAGAGAACATTGAAGCAGACGCGCTTCTTGTCTGTATACAGATATTCGTGATATGC 170
 QY 327 AGCGCGCAAGGACGCTACTCGGGTGAGCTTCGAGCAACGGATCTGTGTTGTGCTATC 386
 DB 171 TAGGTCTCAAGGTAGAGCTCTCTGTACTTTTGGTCCCACTGATATCTTTGCTAGAAC 230
 QY 387 TCTTCAGGACACACCGGAAAGTTTATTCATATAGATGGACACCGGAGAGAACCGAT 446
 DB 231 GCTCAAGGTCTATACCGGAAAGTGTATTCATTGATTTGGACTTCAGAAAAGATAGGAT 290
 QY 447 TCTCAGTCATCTCAAGATGGAGATTAATCTGTGGATGCTCTAAACGAGTCAGAAAAC 506
 DB 291 TGTAGTCATCCCAAGATGGAGATTAATAGTGTGATGCTCTAACAGCAGGAAAC 350
 QY 507 TCATGCTATTAACCTCCCTTGTGCATGGGTATGACATGTGCTTTCTCTCCAAATGGTCA 566
 DB 351 TCATGCTATTAAGCTTCTTGTGCATGGGTATGACATGTGCTTTCTCTCCAAATGGTCA 410
 QY 567 GTCGGTTGCTGTGGATTAAGACAGTATGTTCTATCTTTAGCTTTAGCTTCAACGGC 626
 DB 411 ATCTGTTGCTTGTGGGGCTTTGACAGTGTGTTGCTCTATTTCAATCTTTAATCTCCAC 470
 QY 627 GGACAAGGATGGAACGTGTACCGGTTTCAAGATGTCTACGTGTCACAGGGGATATGTTTC 686
 DB 471 TGATAGGATGGAATCTAAATGTTTTCAGGATGCTTAGTGACATAAAGGTTATGTTTC 530
 QY 687 GTGCTGTCAAGTATGTCCCAATAGGATGCCCACTTATCACCAGTTTCAGGTGATCAAAC 746
 DB 531 ATCTGTGATGATGTTCCAGGTGAAGACACTCACTTAATCACTGTTCTCGAGATCAGAC 590
 QY 747 TTGTATCTTATGGGATGTAACCTACTGTTCTCAAAACTTCTGTTTTTGGCGGTGAATTC 806
 DB 591 ATGTGTTTATGGGATATTAATCTACTGGCTTAGAACATCTGTTTTTGGAGGGAGTTTCA 650
 QY 807 GTCTGGACATCTGTGATGTAAGCTCTCAATCAGTGTGATCAAACTTCAACTGTT 866
 DB 651 GTCTGGACATCTGTGATGTAAGCTTCAATCAGTGTGATCAAACTTCAACTGTT 710
 QY 867 TATATCTGTTTCAATGCGATTCCACAGCAGCTTGTGGGACACTCGTGTGCAAGCCGAGC 926
 DB 711 TGTATCTGTTTCTTGGGATGCGACTGCCAGATTTGGGACACTCGTGTGCAAGCTGAGC 770

Db 672 ATCCGGGACACTGCGAGATGTAACAAAGTGTCTCAATTAGTTATCAAAACCCAGACTGT 731
 QY 867 TATATCTGGTTCATGCGATTCCACAGCAGCGTTGTGGGACACTGCTGTCGACGCGAGC 926
 Db 732 TGTATCTGGTCTCTGTGACACAACTGCTCGACTGTGGGACACCGAGTTGCTAGTCGAGC 791
 QY 927 AGTGGCTACCTTTTCATGGTTCACAGGAGAGTGTAAATAGCGTCAAGTTCTTCCGATGG 986
 Db 792 TCAACGAAACATTTTATGGTTCACAGGAGAGTGTAAATAGCTGTAAAGTTCTTCCCTGATGG 851
 QY 987 GTATAGATTGGGACTGGGATCAGACATGGAACATGCACTGCTGTATGACATAAGGACTG 1046
 Db 852 TAATAGATTGGGAACTGGTTTCAGAGATGGAACCTGCAGATTATTTGACATTAGGACTGA 911
 QY 1047 TCACCAACTCCAGGT---CTATCAGGCCACATGCTGTGATGGTGAACCGGACCTGTGCACCTC 1103
 Db 912 ACACAGCTGCAAGTGTACTACCAAGCGCATGCTGATGGTGTATCCCTCATGTGACTTC 971
 QY 1104 CATTGCAATCTGTGTGACGAGAGACTTTCTTTTCGCTGCTATGCGAGCAACAACTTG 1163
 Db 972 CATGGCAATTTCTATCTCAGGCGCTCTCTCTTTTGTGCGATACTC---AAATGGTGA 1028
 QY 1164 CTACGTTTGGGATACCTCTCTGGGAGAGGTTGTATTGGATTTCGGGATTACAGCAGGATTC 1223
 Db 1029 TTATGTGTGGGACACCTCTATTAGCAAGGTGGTCTTAAACTTGGGAGGAGTTCAAAACTC 1088
 QY 1224 ACACAGGAATPAGAAATAGCTGTTTGGGGTTGTGACGAGATGGAAGTGCAATTTGTACAGG 1283
 Db 1089 TCATGAAGGCGAATAAGTTGCTGGGACTGTGAGCTGATGGAAGCGCTTATGTACAGG 1148
 QY 1284 AAGTTGGGATTCAAATCTAAAGATATGGGCGTTTGGAGGACACAGGAGTGAATTTGA 1341
 Db 1149 AAGTTGGGATACAAACCTGAAGATTGGGCTTTTGGAGGACACAGAGTGTGATCTGA 1206

RESULT 8

ADH78729

ID ADH78729 standard; DNA; 1600 BP.

XX AC

XX ADH78729;

XX XX

XX 22-APR-2004 (first entry)

XX XX

XX Tobacco AGB1 gene ortholog, SEQ ID 9.

XX DE

XX plant; agronomic trait; flowering; fruit yield; seed yield; root biomass;

XX KW

XX seed size; seed shape; stem branch number; plant size; AGB1; gene; de;

XX KW

XX Tobacco; GPA.

XX XX

XX Nicotiana tabacum.

XX OS

XX WO2004003146-A2.

XX XX

XX 08-JAN-2004.

XX XX

XX 24-JUN-2003; 2003WO-US020001.

XX XX

XX 28-JUN-2002; 2002US-0392730P.

XX PR

XX 05-FEB-2003; 2003US-0445208P.

XX PR

XX (PARA-) PARADIGM GENETICS INC.

XX PA

XX (UNYC-) UNIV NORTH CAROLINA.

XX PA

XX Boyes D, Davis K, Jones A, Ullah H, Chen J, Mulpuri R;

XX PI

XX Chatterjee A, Ward MP;

XX XX

XX WPI; 2004-083031/08.

XX DR

XX P-PSDB; ADH78730.

XX XX

PT Altering a plant agronomic trait, useful for improving plant agronomic
 PT traits comprising regenerating plants having stably integrated expression
 PT cassette, where the regenerated plant has an altered agronomic trait.

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

SQ

Disclosure; SEQ ID NO 9; 180pp; English.

The present invention relates to a method for altering a plant agronomic trait selected from time to flowering, duration of flowering, fruit yield, seed yield, root biomass, seed size, seed shape, number of stem branches or size of plant. The method for altering a plant agronomic trait comprises: introducing into a plant cell an expression cassette comprising a nucleotide sequence operably linked to a promoter that is operable within the plant cell, where the nucleotide sequence is selected from: a nucleotide sequence antisense to a plant AGB1 or an AGB1 ortholog (ADH78721-ADH78722 and ADH78725-ADH78746); a nucleotide sequence comprising an inverted repeats of AGB1 or an AGB1 ortholog; a nucleotide sequence encoding a dRNA comprising a first RNA complementary to at least 25 consecutive nucleotides of a plant AGB1 or an AGB1 ortholog and a second RNA substantially complementary to the first RNA; a nucleotide sequence that is AGB1 or an AGB1 ortholog; or a nucleotide sequence that is GPAL or a GPAL ortholog (ADH78723-ADH78724 and ADH78747-ADH78781); and regenerating a plant that has a stably integrated expression cassette from the plant cell, where the regenerated plant has an altered agronomic trait. The method is useful for improving plant agronomic traits including duration of flowering, fruit yield, root biomass, seed size, seed shape, number of stem branches or plant size. The transgenic plants obtained are useful as parents to produce progeny plants and plant lines.

Sequence 1600 BP; 420 A; 343 C; 356 G; 481 T; 0 U; 0 Other;

Query Match 40.2%; Score 654.8; DB 12; Length 1600;

Best Local Similarity 74.7%; Pred. No. 36-192;

Matches 850; Conservative 0; Mismatches 282; Indels 6; Gaps 2;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

Db 742 ATCCGGGCACACTGCAGATGTACAAAGTGTCTCAATTAGTTCATCAAAACCCAGACTGTT 801
QY 867 TATATCTGTTTCATCGATTCACAGCACGGTTGTGGGACACTCGTGTCTGCAAGCCGAGC 926
Db 802 TGTATCTGGTCTCTGTACACAACTGCTGGACTGTGGGACACCCGAGTTGCTAGTCGAGC 861
QY 927 AGTGGTACCTTTCATGTCACAGGAGATGTTAATACGGTCAAGTCTTTCCGGATGG 986
Db 862 TCAACGAACATTTTATGGTTCACAGGGAGATGTTAATCTGTAAGATTTCTCCCTGATGG 921
QY 987 GTATAGATTTGGACTGCACTCAGACATGCAAGATGCGAGCTGTATGACATAAGACTGG 1046
Db 922 TAAATAGATTTGGAATGGTTTCAAGATGGAACCTGCAGATTATTTGACATTTAGACTGG 981
QY 1047 TCACCAACTCCAGGT---CTATCAGCCACATGGTGATGGTGAGAACGGACCTGTCACTTC 1103
Db 982 ACACCACTGCAAGTGTACTACCAAGCCGATGGTGTATGATATCCCTCATGTGACTTC 1041
QY 1104 CATTCGATCTCTGTCTAGGGAGACTTCTTTTCGTGGCTATGGAGCAACACACTTG 1163
Db 1042 CATGGCATTTCTATCTCAGGCCGCTTCTCTCTTTGTCGGATATCTC---AATGGTGAATG 1098
QY 1164 CTACGTTTGGGATACCTCTTCGGAGAGCTTCTATTGGATTGGGATTACAGCAGGATTC 1223
Db 1099 TTATGTGTGGGACACCTTATTAGCAAGGTGGTCTTAACTTGGGAGGATTCAAAACCTC 1158
QY 1224 ACACAGGAATAGATAAGCTGTTTGGGGTTGTGACAGATGGAAGTGCATTTGTGTACAGG 1283
Db 1159 TCATGAAGGGCGAATAAGTTGCTCGGACTGTGAGTGAAGCGCTTATGTACAGG 1218
QY 1284 AGTTTGGGATTCAAATCTAAGATATGGCGTTTGGAGGACACAGAGAGTGAATTGA 1341
Db 1219 AAGTTTGGGATACAAACCTGAAGATTGGGCTTTTGGAGGGGCAAGAGTGTGATCTGA 1276

RESULT 9

ID ADH78733 standard; DNA; 1434 BP.
XX AC ADH78733;
XX DT 22-APR-2004 (first entry)
XX Tobacco AGB1 gene ortholog, SEQ ID 13.
DE Plant; agronomic trait; flowering; fruit yield; seed yield; root biomass;
KW seed size; seed shape; stem branch number; plant size; AGB1; gene; ds;
KW Tobacco; GPA.
XX Nicotiana tabacum.
OS WO2004003146-A2.
XX PN 08-JAN-2004.
XX PD 24-JUN-2003; 2003WO-US020001.
XX PF 28-JUN-2003; 2002US-0392730P.
XX PR 05-FEB-2003; 2003US-0445208P.
XX PA (PARA-) PARADIGM GENETICS INC.
PA (UNNC-) UNIV NORTH CAROLINA.
XX PI Boyes D, Davies K, Jones A, Ullah H, Chen J, Mulpuri R;
PI Chatterjee A, Ward MP;
XX WPI; 2004-083031/08.
DR P-PSDB; ADH78734.
XX Altering a plant agronomic trait, useful for improving plant agronomic
PT traits comprising regenerating plants having stably integrated expression
PT cassette, where the regenerated plant has an altered agronomic trait.
XX

PS Disclosure; SEQ ID NO 13; 180pp; English.
XX The present invention relates to a method for altering a plant agronomic
CC trait selected from time to flowering, duration of flowering, fruit
CC yield, seed yield, root biomass, seed size, seed shape, number of stem
CC branches or size of plant. The method for altering a plant agronomic
CC trait comprises: introducing into a plant cell an expression cassette
CC comprising a nucleotide sequence operably linked to a promoter that is
CC operable within the plant cell, where the nucleotide sequence is selected
CC from: a nucleotide sequence antisense to a plant AGB1 or an AGB1 ortholog
CC (ADH78721-ADH78722 and ADH78725-ADH78746); a nucleotide sequence
CC comprising an inverted repeats of AGB1 or an AGB1 ortholog; a nucleotide
CC sequence encoding a dsRNA comprising a first RNA complementary to at
CC least 25 consecutive nucleotides of a plant AGB1 or an AGB1 ortholog and
CC a second RNA substantially complementary to the first RNA; a nucleotide
CC sequence that is AGB1 or an AGB1 ortholog; or a nucleotide sequence that
CC is GPA1 or a GPA1 ortholog (ADH78723-ADH78724 and ADH78747-ADH78781); and
CC regenerating a plant that has a stably integrated expression cassette
CC from the plant cell, where the regenerated plant has an altered agronomic
CC trait. The method is useful for improving plant agronomic traits
CC including duration of flowering, fruit yield, root biomass, seed size,
CC seed shape, number of stem branches or plant size. The transgenic plants
CC obtained are useful as parents to produce progeny plants and plant lines.
XX Sequence 1434 BP; 395 A; 291 C; 324 G; 424 T; 0 U; 0 Other;

Query Match 40.0%; Score 651; DB 12; Length 1434;
Best Local Similarity 74.5%; Pred. No. 4.3e-191;
Matches 848; Conservative 0; Mismatches 285; Indels 6; Gaps 2;

QY 207 AATGTCTGTCCTCGAGCTCAAAGAACGCCAGCCGTCGCTACGGAGACCGTTAATAACCT 266
Db 74 AATGTCACTGAAGAGCTGAAAGAGCGGCATATGCGCGCTACACAAACTGTAATATGATCT 133
QY 267 CCGTGACCGACTAGACAGAGACGCCCTCCAGCTCTCGATACCCGATCTGGTTTGTTCGTAC 326
Db 134 CCGTGAAACCTTAAGCAGAAAGCGTCTCCAAATTAATCGACACTGATGTATCTGGGTATGC 193
QY 327 AGCGGCGCAAGGACGTACTCGGGTGAGCTTCGGAGCAACGGATCTGGTTTGTTCGTAC 386
Db 194 AAGTGCAGAGTAAACTCCGGTCATCTTTGGCCCAACAGATCTGGTTTGTTCGTAGGAT 253
QY 387 TCTTCAGGACACACCGGAAAGGTTTATTCATAGATTGGACACGGAGAGAACCGGAT 446
Db 254 CCTGCAAGGACACACTGGAAGGTTATTTCACTGGATTGGACTCCAGAAAAGAAATCGTAT 313
QY 447 TGTCAGTGCATCTCAAGATGGAGATTAATCGTGTGGAATGCTCTAAGAGCTCAGAAAC 506
Db 314 AGTCAGTGCATCCCAAGATGGCAGATTAAATAGTGTGAATGCTCTCAACCCAGAAAC 373
QY 507 TCATGCTATTAAACTCCCTTGTGCATGGGTTATGACATGTGCTTTCTCTCCAAATGGTCA 566
Db 374 CCATGCAATTAAAGCTTCCATGTGCTTGGTTATGACCTGCGCCTTCTCTCTAGTGGCA 433
QY 567 GTCGGTTGCGTGTGGTGGATTAGACAGGTATGTTTCTATCTTTAGCCCTTAGCTCAACCGC 626
Db 434 GTCTGTTGCTGCGTGGCCTTGACAGTGTCTGCTCTATCTTCAACTTAATATTCACCGAT 493
QY 627 GGACAAGGATGGAACTGTACCGGTTTCAAGAAATGCTCACTGGTCACAGGGGATATGTTTC 686
Db 494 CGATAAGGATGGGAACCATCTCTGTATCAAGAAATGCTTAGTGGGCATTAAGGGGTATGTGTC 553
QY 687 GTGCTGTCAGTATGTCCCAATAGGATGCCACCTTATCACCAGTTCAAGTGTGATCAAAAC 746
Db 554 TTCTGTCAGTATGTTCAGATGAGGATACCTACGTAATAACTAGTCTTGTGTGATCAAAAC 613
QY 747 TTGTATCTTATGGGATGTAACACTGGTCTCAAAACTTCTGTTTTTGGCGGTGAATTCA 806
Db 614 ATGTGTCCTTTGGGATATAACTACTGGCTTAAGAACTTCTGCTTTTGGAGGTGAGTTTCA 673
QY 807 GTCTGGACATACCTCTGATGTACTAAGCTCTCAATCAGTGGATCAAAACCCAACTGGTT 866
Db 674 ATCCGGGCACACCCGACATGTACAAAGTGTCTCAATTAGTTTCATCAAAACCCACACTGTT 733

QY 867 TATATCTGTTTCATGCGATCCACAGCAGCGTTGTGGGACACTCTGCTGTCGAAGCCGAGC 926
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 734 TGTGTCGGGCTCTGTGACTCAACTGCTCGACTATGGGACACCCGAGTTGCTAGTCGAGC 793
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 927 AGTGGCTACCTTTTCATGGTCACAGGAGAGATGTTAATACGGTCAAGTTCTTTCCGATGG 986
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 794 TCACGACATTTTATGGTCATGAGGAGATGTTAATCTGTAAAGTTCTTCCCTGATGG 853
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 987 GTATAGATTGGGACTGGATCAGACGATGGAACATGCGAGCTGTATGACATAGGACTGG 1046
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 854 TAATAGATTGGACTGGTTTCAGATGATGAACCTGCAGATTATTGACATTAGGACTGG 913
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1047 TCACCACTCCAGGT---CTATCAGCCACATGCTGATGTGAGAGCGACCTGTCACTTC 1103
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 914 ACACCACTGCAAGTGTACTACACGCGCATGCTGATGATGATGATGATGATGATGATGATG 973
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1104 CATTCATCTCTGTCAGGAGACTCTTTTCGCTGCTATGCGAGCAACACACTTG 1163
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 974 CATGGCATTTTCTATCTCAGGCCGCTCTCTCTTTGCGGGTACTC---AAATGGTGATG 1030
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1164 CTACGTTTGGGATPACCTCTTGGGAGAGGTGTTGTTGTTGGATTTACAGCAGGATTC 1223
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1031 TTATGTCGGACACCTTATTAGCAAGGTGGTCTTAACCTTGGGAGCAGTTCAAACTC 1090
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1224 ACACGGAATAGAAATAGCTGTTTGGGTTGTCAGCAGATGGAAGTGCATTGTGTACAGG 1283
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1091 TCATGAAGGCGGAATAGTTGCTGGGACTGTGATGATGATGATGATGATGATGATGATGATG 1150
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1284 AGTTGGGATCAAACTAAGATATGGCGTTTGGAGGACACACAGAGATGATTTGAA 1342
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 1151 AAGTTGGGATACAACTGAAGATTTGGGCTTTTGGAGGGACACAGAGTGTGATCTGAA 1209

RESULT 10

AE26929 ID AEB26929 standard; cDNA; 1676 BP.
 XX AC AEB26929;
 XX DT 22-SEP-2005 (first entry)
 XX DE Pinus radiata cell cycle protein cDNA SEQ ID NO 90.
 XX KW plant protectant; fungicide; plant growth regulant; gene therapy;
 KW cell cycle; gene expression; plant; transgenic plant; microarray; gene;
 KW ss; wood.
 XX OS Pinus radiata.
 XX PN WO2005065339-A2.
 XX PD 21-JUL-2005.
 XX PF 30-DEC-2004; 2004WO-US043804.
 XX PR 30-DEC-2003; 2003US-0533036P.
 XX PA (ARBO-) ARBOGEN LLC.
 XX PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX PI Forster RL, Connett MB, Emerson SJ, Grigor MR, Higgins CM;
 PI Lund ST, Magusin A, Kodrzycki RJ;
 XX WPI: 2005-506765/51.
 XX P-PSDB; AEB27189.
 XX PT New polynucleotide encoding a plant cell cycle protein, useful for
 XX modifying plant development and altering plant phenotype.
 XX PS Claim 1; SEQ ID NO 90; 499pp; English.
 XX CC The invention describes an isolated polynucleotide comprising: (i) a

CC sequence of SEQ ID NOS: 1-237 or their conservative variants; (ii) a
 CC sequence encoding the catalytic or substrate-binding domain of a
 CC polypeptide of SEQ ID NOS: 261-497, where the polynucleotide encodes a
 CC polypeptide having the activity of the polypeptide of SEQ ID NOS: 261-497
 CC ; or (iii) a nucleic acid sequence of SEQ ID NOS: 471-697. Also described
 CC are: (1) a DNA construct comprising at least one polynucleotide having
 CC the sequence of SEQ ID NOS: 1-237 or their conservative variants; (2) a
 CC plant cell transformed with the DNA construct of (1); (3) a transgenic
 CC plant comprising the plant cell of (2); (4) a method of making a
 CC tree which has been transformed with the DNA construct of (1); (6) a
 CC method of making wood or wood pulp; (7) an isolated polypeptide
 CC comprising an amino acid sequence encoded by the new isolated
 CC polynucleotide or comprising any of the amino acid sequences of SEQ ID
 CC NOS: 261-497; (8) a method of altering a plant phenotype of a plant; (9)
 CC a method of correlating gene expression in two different samples; (10) a
 CC method of correlating the possession of a plant phenotype to the level of
 CC gene expression in the plant of one or more genes; (11) a method of
 CC correlating gene expression to a stage of the cell cycle; (12) a
 CC combination, for detecting expression of one or more genes, comprising
 CC two or more oligonucleotides, where each oligonucleotide is capable of
 CC hybridizing to a nucleic acid sequence of SEQ ID NOS: 1-237 or to gene
 CC product encoded by a nucleic acid sequence of SEQ ID NOS: 1-237; (13) a
 CC microarray comprising the combination of (12) provided on a solid
 CC support, where each of the two or more oligonucleotides occupies a unique
 CC location on the solid support; (14) a method for detecting one or more
 CC genes in a sample; (15) a method for detecting one or more nucleic acid
 CC sequences encoded by one or more genes in a sample; and (16) a kit, for
 CC detecting gene expression, comprising the microarray of (13) together
 CC with one or more buffers or reagents for a nucleotide hybridization
 CC reaction. The polynucleotides polypeptides, DNA construct, composition,
 CC and methods are useful for modifying plant development and altering plant
 CC phenotype. This sequence encodes a cell cycle protein isolated in the
 CC invention.
 XX SQ Sequence 1676 BP; 387 A; 407 C; 437 G; 445 T; 0 U; 0 Other;

QY 208 ATGTCTGTCTCGAGCTCAAGAAGCCGAGCGCTGCTACGGAGACCGTTAATACCTC 267
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 257 ATGTCTGTCTCGAGCTCAAGAAGCCGAGCGCTGCTACGGAGACCGTTAATACCTC 316
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 268 CGTGACCACTTTAGACAGAGAGCGCTCCAGCTCTCTCGATACCGATGTGCGAGGATTCA 327
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 317 CGGGAGCGCTCAAGCAGAAGCGAGTGCAGCTGCTCGACACCGAGCTGCGGAGTACGG 376
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 328 GCGGCGCAAGGAGCTACTCGGGTGTGAGCTTCGGAGCAACGGATCTGGTTTGTGTCTACT 387
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 377 AGGACTCTAGGGGAAAGTCTCCGCTCACTTTTGGGGGAGCGGATCTGGTGTGCGAGGACC 436
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 388 CTTTCAGGAGACACACCGGAAAGTTTATTATTAGATTGGACACCGAGAGGAAACCGGATT 447
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 437 TTGCAAGGCGCACACTGGGCAAGGTGTACTCCTAGATTGGACTCTCTGAAAGGAATGCCATT 496
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 448 GTCAAGTGTCTCAAGATGGGAGATTAAATCGTGTGGAATGCTCTTAACCGAGTCAAGAACT 507
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 497 GTCAAGTGTCTCAAGATGGGAGATTAAATCGTGTGGAATGCTCTTAACCGAGTCAAGAACT 556
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 508 CATGTCTATTAACTCTCCCTTGTGATGGGTTATGACATGTGCTTCTCTCCAAATGTGTAG 567
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 557 CATGCCATTAAGGCTCTCTGTGATGGGTAATGACATGTGCTTTTGCACCGAATGTGTAG 616
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 568 TCGGTTGCGTGTGGATTAGACAGTGTATCTTCTATCTTTAGCTTACGCTCAACGGCG 627
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 617 TCTGTGCTGTGTGGCGGTTTTCACAGCGTATGCTCTATCTTAACCTCAATTTCCCGGTT 676
 QY | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 628 GACAGGATGGAACCTGTACCGGTTTCAAGAAATGCTCCTCAGTGTGACAGGGGATATGTTTCG 687
 Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
 677 GACAGGATGGAACCTGTACCGGTTTCAAGAAATGCTCCTCAGTGTGACAGGGGATATGTTTCG 736

QY 688 TCCTGTCTAGTATGTCCTCCAAATGAGGATGCCACCTTATCACCAGTTCAGGTGATCAAACT 747
 Db |||||
 QY 737 TCATGTCTAGTATGTCACGACGAGATGCTACCTGATACAGGATCTGGTACCAACA 796
 Db |||||
 QY 748 TGTATCTTATGGATGTAACCTACTCTCAAACTCTGTTTGGCGGTGAATTTTCA 807
 Db |||||
 QY 797 TGTGTCTTGTGGATATACCAAGGTCTGAGAACTTCTGTCTTTGGAGGAAATTTCAA 856
 Db |||||
 QY 808 TCTGACATCTCTGTGATGTACTAAGCGCTCTCAATCAGTGGATCAAAACCAAACTGGTTT 867
 Db |||||
 QY 857 TCAGGGCACCTGCTGATGTTTTAAGTGTGTCATCAATGGATCGAGCCCAAGATATTT 916
 Db |||||
 QY 868 ATATCTGTTTCATGATTCACAGACCGTTGTGGGACATCTGTGCTGCAAGCCGAGCA 927
 Db |||||
 QY 917 GTTCTGTTTCAATGATTAACCTGCTAGGATGTGGGATCTGCTGTGTCGAAGTCGAGCC 976
 Db |||||
 QY 928 GTGCGTACTCTTTCAGTCTCAGCAGGAGATGTTAATACGTCACGTTCTTTCCGATGGG 987
 Db |||||
 QY 977 GTTCATACATATCATGGACATGAGCGGATGTAATGCCGTAAAGTCTTTTCCAGATGGA 1036
 Db |||||
 QY 988 TATAGATTTGGACTGGATCAGACGATGGAACATGCAGGCTGTATCACAATAGGACTGTG 1047
 Db |||||
 QY 1037 AATAGATTTGGAACCTGGTCTGATGACGCTACTTGCAGGCTCTTTGACATCAGACAGG 1096
 Db |||||
 QY 1048 CACCAACTCCAGGT---CTATCAGGCACATGATGATGTGAGAACGACCTGTCTACCTCC 1104
 Db |||||
 QY 1097 CATGAACCTTCAAGTGTACTATCAACAGCGTGGCATCGATGAGATCCCAATGTCACTCC 1156
 Db |||||
 QY 1105 ATTGCATCTCTGTCTCAGGAGACTCTTTTCGCTGGCTATGCGGACGACCAACACTTGC 1164
 Db |||||
 QY 1157 ATTGCATTTTCATTTCCGGGAGGCTGCTAAATGCTGGATATCAAAATGGTAC---TGC 1213
 Db |||||
 QY 1165 TACGTTTGGGATPACCTCTTGGGAGAGGTTGATTTGGATTTGGGATTAACAGCAGGATTC 1224
 Db |||||
 QY 1214 TTTGTGGGATACATTAATGCCCCAGGTTGTTGAACTTGGGATCGTTGAGAACTCA 1273
 Db |||||
 QY 1225 CACAGGAATAGAAATGCTGTTTGGGTTGTGACAGATGGAAGTGCAATGTTGTACAGGA 1284
 Db |||||
 QY 1274 CACGAAGTTCGGATCAGCTGTTTGGGTGTCTGCTGATGGAAGTGCCCTTGTGTACTGCG 1333
 Db |||||
 QY 1285 AGTTGGGATTCAAATCTAAAGATATGGGCTTTGGAGGACACAGCAGAGATGATTTGA 1341
 Db |||||
 QY 1334 AGTTGGGATACAAACCTAAAGATTTGGGCTTTTGGAGGATTCGAGGGTGACTTAA 1390
 Db |||||

RESULT 11
 ID ADH78725
 XX ADH78725 standard; DNA; 1540 BP.
 AC ADH78725;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Potato AGB1 gene ortholog, SEQ ID 5.
 XX
 KW Plant; agronomic trait; flowering; fruit yield; seed yield; root biomass;
 KW seed size; seed shape; stem branch number; plant size; AGB1; gene; ds;
 KW Potato; GPA.
 XX
 KW Solanum tuberosum.
 OS
 XX
 PN WO2004003146-A2.
 XX
 PD 08-JAN-2004.
 XX
 PF 24-JUN-2003; 2003WO-US020001.
 XX
 PR 28-JUN-2002; 2002US-0392730P.
 PR 05-FEB-2003; 2003US-0445208P.
 XX
 PA (PARA-) PARADIGM GENETICS INC.
 PA (UYNC-) UNIV NORTH CAROLINA.
 XX

PI Boys D, Davis K, Jones A, Ullah H, Chen J, Mulpuri R;
 PI Chatterjee A, Ward MP;
 XX
 DR WPI; 2004-083031/08.
 DR P-PSDB; ADH78726.
 XX
 PT Altering a plant agronomic trait, useful for improving plant agronomic
 PT traits comprising regenerating plants having stably integrated expression
 PT cassette, where the regenerated plant has an altered agronomic trait.
 XX
 PS Disclosure; SEQ ID NO 5; 180pp; English.
 XX
 CC The present invention relates to a method for altering a plant agronomic
 CC trait selected from time to flowering, duration of flowering, fruit
 CC yield, seed yield, root biomass, seed size, seed shape, number of stem
 CC branches or size of plant. The method for altering a plant agronomic
 CC trait comprises: introducing into a plant cell an expression cassette
 CC comprising a nucleotide sequence operably linked to a promoter that is
 CC operable within the plant cell, where the nucleotide sequence is selected
 CC from: a nucleotide sequence antisense to a plant AGB1 or an AGB1 ortholog
 CC (ADH78721-ADH78722 and ADH78725-ADH78746); a nucleotide sequence
 CC comprising an inverted repeats of AGB1 or an AGB1 ortholog; a nucleotide
 CC sequence encoding a dsRNA comprising a first RNA complementary to at
 CC least 25 consecutive nucleotides of a plant AGB1 or an AGB1 ortholog and
 CC a second RNA substantially complementary to the first RNA; a nucleotide
 CC sequence that is AGB1 or an AGB1 ortholog; or a nucleotide sequence that
 CC is GPAl or a GPAl ortholog (ADH78723-ADH78724 and ADH78747-ADH78781); and
 CC regenerating a plant that has a stably integrated expression cassette
 CC from the plant cell, where the regenerated plant has an altered agronomic
 CC trait. The method is useful for improving plant agronomic traits
 CC including duration of flowering, fruit yield, root biomass, seed size,
 CC seed shape, number of stem branches or plant size. The transgenic plants
 CC obtained are useful as parents to produce progeny plants and plant lines.
 XX
 SQ Sequence 1540 BP; 423 A; 314 C; 331 G; 472 T; 0 U; 0 Other;

Query Match 39.4%; Score 642.4; DB 12; Length 1540;
 Best Local Similarity 73.9%; Pred. No. 2.1e-188;
 Matches 843; Conservative 0; Mismatches 291; Indels 6; Gaps 2;
 QY 205 GGATGTCTGTCTCGAGCTCAAGAACGCCACCGCTCGCTACGAGACCGTTAATAAC 264
 Db |||||
 QY 93 GAAATGTCTGAGTTCGGAGCTGAAAGCGGACATGCGCCGCTACACAGACTGTAATGAT 152
 Db |||||
 QY 265 CTCGTGACAGCTTAGACAGACAGCGCTCCAGCTCTCCGATACCGATGTGCCAGGAT 324
 Db |||||
 QY 153 CTCGTGAAAACCTTAAGCAGACAGCGTCTCAATTAATCTCGACAGAGATGTTCTGGGTAT 212
 Db |||||
 QY 325 TCAGCGCGCAAGGACGCTACTCTCGGTGAGCTTCGGAGCAACCGATCTGTTTGTCTGT 384
 Db |||||
 QY 213 GCAAAGACGCAAGGTAAACCTCCGTAACGTTCCGCCCAACAGATCTAGTTTGTGTAGG 272
 Db |||||
 QY 385 ACTCTTCAGGACACACCGGAAAGTTTATTCATAGATTGGACACCGGAGAGGACCGG 444
 Db |||||
 QY 273 ATCTGCAAGGACACACTGGAAAGGTCTATTTCACTGGACTGGACTCTCTGAAAAAATCGT 332
 Db |||||
 QY 445 ATTGTCTAGTCTCAAGATGGGAGATTAATCGTGTGGAATGCTTAACGAGTCAGAAA 504
 Db |||||
 QY 333 ATAGTCTAGTGCATCCCAAGATGGTAGATTAATAGTGTGAATGCTCTCAAGACCGAGAA 392
 Db |||||
 QY 505 ACTCATGCTATTAAACTCCCTTGTGCATGGGTATAGCATGTGTTCTCTCCAAATGGT 564
 Db |||||
 QY 393 ACCATGCAATTAAGCTTCCATGTGTTGGTATGACCTGTGCTCTCTCTCTAGTGA 452
 Db |||||
 QY 565 CAGTCTGTTGCTGTGGATTAGACAGTGTATGTTCTATCTTTAGCTTAGCTCAACG 624
 Db |||||
 QY 453 CAGTCTGTTGCTGTGGCGGCTTTGACAGTGCCTGCTATCTTTCAACTTAAATTCACCA 512
 Db |||||
 QY 625 GCGGACAGGATGGAACCTGTACCGGTTTCAAGAAATGCTCACTGTGTACAGGGGATATGTT 684
 Db |||||
 QY 513 ATTGATAGGATGGATCCATCCAGTATCGAATGCTTAGTGGGCATAGGGGTATGTG 572
 Db |||||
 QY 685 TCGTGTGTCTAGTATGTCCTCAAAATGAGGATGCCACCTTATTCACAGGTTTCAGGTGATCAA 744
 Db |||||

Db 573 TCTTCGTGTCAGTATGTTCCGGATGAGGATATCTCACCTAAATAACTAGTCTCGGTGATCAA 632
Qy 745 ACTTGTATCTTATGGATGTAATCTGCTCTCAAAACCTTCTGTTTGGCGGTGAATTT 804
Db 633 ACATGTGTAATTTGGATATACTACTGCGCTTAAGACTTCTGTGTTGGAGGTGAGTTT 692
Qy 805 CAGTCTGGACATCTGCTGATGTAATAAGCGTCTCAATCAGTGGATCAAAACCAACTGG 864
Db 693 CAATCTGGGCACATCTGATGATGTAATCAAGTCTCAATAGTTCATCTAAACCCCAACTA 752
Qy 865 TTATATCTGTTCTGATGATTTCCACAGCAGCGTGTGGGACACTGCTGCTCAAGCCGA 924
Db 753 TTGTGTCCTGGGTCCTGTGACAACTGCTGACGTGGGACACCGAGTGTCTAGTGA 812
Qy 925 GCAGTCTGCTACCTTTTCAATGTCACGAGGAGATGTTAATACGGTCAAGTCTTCTCGGAT 984
Db 813 GCTCAACGAACATTTTCATGACACGAGAGTGATTTACTGTAAGTCTTCTCCCTGAC 872
Qy 985 GGTATATAGATTTGGGACTGGATCAGACGATGGAACATGACAGGTGTATGACATAGGACT 1044
Db 873 GGTAAATAGATTTGGAACTGTTTCAGATGATGCGAGCTGCAGATTAATTTGACATTAGGACT 932
Qy 1045 GGTCAACCACTCCAGCT---CTATCAGCCACATGGTGTGAGTGTGAGAACGACTGTCAAC 1101
Db 933 GGACACAGCTGCAAGTATATACCAACCCGATGCTGACGCTGACATCCCTCATGTGACT 992
Qy 1102 TCCATTTGCATTTCTGTCAGGAGACTTCTTTTCGCTGGCTATCCGAGCAACACACT 1161
Db 993 TCCATTTGCATTTCTATCTCAGCCGCTCTCTTTGTCGGGTACTC---TAATGTGAT 1049
Qy 1162 TGCTACGTTTGGGATACCTCTTTGGGAGAGTGTGATTTGGATTTGGATTTACAGCAGGAT 1221
Db 1050 TGTTACGTTTGGGACACCCCTATTAGCAAGGTGCTCTAAACTTAGGATCAGTTCAAAAC 1109
Qy 1222 TCACACAGGATAGATAGCTGTTTGGGTTGTCAGCAGATGGAAGTGCATTTGTGTACA 1281
Db 1110 TCTCATGAAGGCGGATTAAGTGTCTGCGGACTGTCTAGCTGATGGAAGTGCCTTATGTACA 1169
Qy 1282 GGAAGTTGGGATTCAAATCTAAAGATATGGCGCTTTTGGAGGACACAGGAGAGTGTGAT 1341
Db 1170 GGAAGTTGGGATCAAACTGATGATTTGGGCTTTTGGAGGACACAGAGTGTGATCTGA 1229

RESULT 12
ADH78731

ID ADH78731 standard; DNA: 1560 BP.
XX AC ADH78731;
XX DT 22-APR-2004 (first entry)
XX DE Tobacco AGB1 gene ortholog, SEQ ID 11.
XX KW Plant; agronomic trait; flowering; fruit yield; seed yield; root biomass;
KW seed size; seed shape; stem branch number; plant size; AGB1; gene; da;
KW Tobacco; GPA.
XX OS Nicotiana tabacum.
XX PN WO2004003146-A2.
XX PD 08-JAN-2004.
XX PF 24-JUN-2003; 2003WO-US020001.
XX PR 28-JUN-2002; 2002US-0392730P.
XX PR 05-FEB-2003; 2003US-0445208P.
XX PA (PARA-) PARADIGM GENETICS INC.
XX PA (UYNC-) UNIV NORTH CAROLINA.
XX PI Boyes D, Davis K, Jones A, Ullah H, Chen J, Mulpuri R;

PI Chatterjee A, Ward MP;
XX WPI; 2004-083031/08.
DR P-PSDB; ADH78732.
XX
PT Altering a plant agronomic trait, useful for improving plant agronomic
PT traits comprising regenerating plants having stably integrated expression
PT cassette, where the regenerated plant has an altered agronomic trait.
XX
PS Disclosure; SEQ ID NO 11; 180pp; English.
XX
CC The present invention relates to a method for altering a plant agronomic
CC trait selected from time to flowering, duration of flowering, fruit
CC yield, seed yield, root biomass, seed size, seed shape, number of stem
CC branches or size of plant. The method for altering a plant agronomic
CC trait comprises: introducing into a plant cell an expression cassette
CC comprising a nucleotide sequence operably linked to a promoter that is
CC operable within the plant cell, where the nucleotide sequence is selected
CC from: a nucleotide sequence antisense to a plant AGB1 or an AGB1 ortholog
CC (ADH78721-ADH78722 and ADH78725-ADH78746); a nucleotide sequence
CC comprising an inverted repeats of AGB1 or an AGB1 ortholog; a nucleotide
CC sequence encoding a dsRNA comprising a first RNA complementary to at
CC least 25 consecutive nucleotides of a plant AGB1 or an AGB1 ortholog and
CC a second RNA substantially complementary to the first RNA; a nucleotide
CC sequence that is AGB1 or an AGB1 ortholog; or a nucleotide sequence that
CC is GPAL or a GPAL ortholog (ADH78723-ADH78724 and ADH78747-ADH78781); and
CC regenerating a plant that has a stably integrated expression cassette
CC from the plant cell, where the regenerated plant has an altered agronomic
CC trait. The method is useful for improving plant agronomic traits
CC including duration of flowering, fruit yield, root biomass, seed size,
CC seed shape, number of stem branches or plant size. The transgenic plants
CC obtained are useful as parents to produce progeny plants and plant lines.
XX
SQ Sequence 1560 BP; 395 A; 317 C; 338 G; 510 T; 0 U; 0 Other;

Query Match 39.4%; Score 641.2; DB 12; Length 1560;
Best Local Similarity 74.5%; Pred. No. 5.1e-188;
Matches 848; Conservative 0; Mismatches 283; Indels 7; Gaps 3;

Qy 207 AATGTCTGTCTCCGAGCTCAAGAACGCCACGCCCTCGCTACGGAGACCGTTAATAACCT 266
Db 119 AATGTCACTGACAGAGCTGAAAGAGCGGCATATGCCGCTACACAGACTGTAACTGATCT 178
Qy 267 CCGTGACACGCTTACAGAGACCGCTCCAGCTCTCGATACCGATGTGGCAGGATTC 326
Db 179 CCGTGAAAACCTTAAGCAGAACCGCTCCCAATTACTCGACACTGATGTTCTTGGATATGC 238
Qy 327 AGCGGCGCAAGGACGCTACTCGGTGAGCTTCGGAGCAACCGGATCTGTTGTCTGTAC 386
Db 239 AAGTCCGCAAGGTAAACTCCCGTCACTTTGGCCCAACAGATCTGTTGTGTAGGAT 298
Qy 387 TCTTCAGGACACACCGGAAAGTGTATTTCATTAGATTGGACACCGGAGAACCGGAT 446
Db 299 CTTGCAAGGACACACTTGAAGGTATATTCACTGGATTGCACTCCAGAAAAGAAATCGTAT 358
Qy 447 TGTCACTGATCTCAAGATGGAGATTAATCTGTGGAATGCTCTTAACAGTCAAGAAAC 506
Db 359 AGTCAGTGCATCCCAAGATGGCAGATTAATAGTGTGGAATGCTCTCACAAGCCAGAAAC 418
Qy 507 TCATGCTATTAAACTCCCTTGTGATGGGTATGACATGTGCTTTCTCTCCAAATGGTCA 566
Db 419 CCATGCAATTAAGCTTCCGTTGCTGGGTATGACCTGGCCCTTCTCTAGTGGGCA 478
Qy 567 GTCGGTTGGCTGTGGATTAGACAGTGTATGTTCTATCTTTAGCCTTAGCTCAACCGC 636
Db 479 GTCTGTTGCTTCGGTGGCGCTTGACAGTGTCTGCTCTATCTAACTTAAATTCGCAAT 538
Qy 627 GGAACAAGGATGGAACCTGTACCGGTTTCAAGAAATGCTCACTGTGTCAAGGGGATATGTTTC 686
Db 539 CGATAAGGATGGGAACCATCTCTGTATCAAGAAATGCTTAGTGGGCATTAAGGGTTATGTGTC 598
Qy 687 GTGCTGTAGTATGTCCTCCAAATGAGGATGCCACCTTATCACCAGTTTCAGGTGATCAAAC 746

Qy	734	CAGGTGATCAACACTGTTATCTTATGGGATGTAACACTACTGGTCTCAAAACTCTGTTTTTG	793	PR	01-APR-1999;	99US-0127462P.
Dd	670	CTGGAGATCAGACATGTGTGTTTATGGGATTAATCTACTGGCCCTTAGAACATCTGCTCTTT	729	PR	06-APR-1999;	99US-0128234P.
Qy	794	GGGGTGAATTTCAAGTCTGGACATACCTGCTGATGTAAGCGTCTCAATCACTGATGATCAA	853	PR	16-APR-1999;	99US-0129845P.
Dd	730	TAGGGAGTTTCAGTCTGGACATCTGCAGATGTAATTTAGCAATTTCCATTAATGATCA	789	PR	19-APR-1999;	99US-0130077P.
Qy	854	ACCCAAACTGGTTTATATCTGGTTTCATCGATTCACACGACGGTGTGGGACACTCGTG	913	PR	21-APR-1999;	99US-0130449P.
Dd	790	ACTCCAAATTTGTTGATCTGGTTCTTGCGATGCGACTGCCAGATTTGTGGGACACTCGTG	849	PR	23-APR-1999;	99US-0130510P.
Qy	914	CTGCAAGCGGACGACGTGCTACTCTTTCATGCTCAGGAGGAGATTTAATACGGTCAAGT	973	PR	28-APR-1999;	99US-0130891P.
Dd	850	TGGCAAGTCGAGCAGTCGGGACATTTTCACGGCCACGAGGAGATGTTAATCTGTCAAA	909	PR	30-APR-1999;	99US-0131449P.
Qy	974	TCTTTCCGGATGGGTATAGATTTGGGACTTGGATCAGACATGGAACATGCGAGCTGTATG	1033	PR	30-APR-1999;	99US-0132048P.
Dd	910	TCTTTCTGTGTAAGAAATAGATTTGGAATCTGGCTCAGAGATGGAATCTGCAGATTTATG	969	PR	04-MAY-1999;	99US-0132484P.
Qy	1034	ACATAAGACTGGTCCACCAACTCCAGGTCT---ATCAGCCACATGGTGTGATGGTGAACG	1090	PR	05-MAY-1999;	99US-0132486P.
Dd	970	ACATTAGGACCGGACACCAACTTCAAGTATATAATCAGCAACACCAACGAAATGG	1029	PR	06-MAY-1999;	99US-0132487P.
Qy	1091	GACCTGTCACTCCGATTCGATCTCTGTGTCAGGAGACTCTTTTCGCTGCTATGCGA	1150	PR	07-MAY-1999;	99US-0132863P.
Dd	1030	CACATGTGACGTCCTCAATTCATTTCCATATCCGGAAGACTTCTTATTCGTGGCTAT---A	1086	PR	11-MAY-1999;	99US-0134218P.
Qy	1151	GCAACAACTTGTCTACGTTTGGGATACCTCTTCGGGAGAGTTGCTATTGGATTGGGAT	1210	PR	14-MAY-1999;	99US-0134219P.
Dd	1087	CAATGGTGTATGCTATGTTGGGATATCTTATTCGCTAAGTGTCTTGAATCTAGAT	1146	PR	14-MAY-1999;	99US-0134219P.
Qy	1211	TACAGCAGATTACACAGGAATAGAATAAGCTGTTTGGGTTGTGACGAGATGGAAGTG	1270	PR	14-MAY-1999;	99US-0134370P.
Dd	1147	CTCTTCAAACTCTCATGAGGCGAGATCACCTGTTTGGGTATGCTGCTGATGGAACG	1206	PR	18-MAY-1999;	99US-0134768P.
Qy	1271	CATTGTGTACAGGAGTTGGGATTCAAATCTAAAGATATGGCGTTTGGAGGACACAGCA	1330	PR	20-MAY-1999;	99US-0135124P.
Dd	1207	CTTATGTACAGGAGTTGGGACACAAATTTAAAGATATGGGCATTTGGAGGCAATGGA	1266	PR	21-MAY-1999;	99US-0135353P.
Qy	1331	GAGTGATTGA 1341		PR	24-MAY-1999;	99US-0135629P.
Dd	1267	AGGTGATTGA 1277		PR	27-MAY-1999;	99US-0136021P.
RESULT 15						
AAC44600						
ID	AAC44600 standard; DNA; 693 BP.					
XX	AAC44600;					
AC	AAC44600;					
DT	18-OCT-2000 (first entry)					
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 43426.					
XX	Arabidopsis thaliana.					
KW	Hybridisation assay; genetic mapping; gene expression control;					
KW	protein identification; signal transduction pathway; metabolic pathway;					
KW	promoter; termination sequence; ss.					
XX	Arabidopsis thaliana.					
OS	Arabidopsis thaliana.					
PN	EP1033405-A2.					
XX	06-SEP-2000.					
PD	25-FEB-2000; 2000EP-00301439.					
PF	25-FEB-1999; 99US-0121825P.					
XX	05-MAR-1999; 99US-012180P.					
PR	09-MAR-1999; 99US-0123548P.					
PR	23-MAR-1999; 99US-0125789P.					
PR	25-MAR-1999; 99US-0126264P.					
PR	29-MAR-1999; 99US-0126785P.					

Qy	766	ACTACTGGTCTCAAAACTTCTG-TTTTGGCGGTGAATTTTCAGTCT-GGACATACTGCTG	823
Db	586	ACTACTGGTCTCAAAACTTCTGTTTTTGGCGGGGAATTTTCAGTCTGGGACATACTGCTG	645
Qy	824	ATGTACTAAGCGTCTCAATCAGTGGATCAAAACCCAACTGGTTTATAT	871
Db	646	ATGTACTAAGCGTCTCAATCAGTGGATCAAAACCCAACTGGTTTATAT	693

Search completed: December 28, 2005, 15:33:31
Job time : 949 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 13:54:02 ; Search time 5978 Seconds

(without alignments)
12749.441 Million cell updates/sec

Title: US-10-602-898A-1

Perfect score: 1629
Sequence: 1 cctgcgtagcagctgttg.....cagaagataaaacgctaacta 1629

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1499.2	92.0	1537	4	CNS0A3HU
2	1289	79.1	1537	4	CNS0A3QO
3	1251.4	76.8	1765	4	CNS0A2EA
4	748.6	46.0	896	2	BE037621
5	565	34.7	663	1	AV823429
6	554.4	34.0	590	1	AV784390
7	552.2	33.9	1685	4	AY109568
8	532	32.7	566	6	CA781376
9	522	32.0	978	7	CK248467
10	516.2	31.7	858	7	CO072720
11	515.4	31.6	871	7	CO072054
12	513.2	31.5	535	1	A1994586
13	511.4	31.4	838	7	CO072562
14	506.6	31.1	645	1	AM058772
15	504.6	31.0	778	8	DN620362
16	504.2	31.0	829	7	CO125466
17	487.8	29.9	502	1	AV560034
18	479	29.4	729	6	CF417476
19	477.2	29.3	783	8	CK050448
20	476.8	29.3	804	2	BG647926
21	470.2	28.9	879	6	CF203907
22	466.4	28.6	783	3	BI422788

23	466	28.6	790	3	BM406213
24	464	28.5	967	7	CK298648
25	459	28.2	885	8	DR912730
26	450.2	27.6	859	7	CK279376
27	449	27.6	461	6	CD528998
28	444.2	27.3	740	6	CD835752
29	442.4	27.2	714	6	CD835721
30	437.6	26.9	713	8	DN958590
31	435.4	26.7	663	7	CN519259
32	435.4	26.7	870	8	DR925535
33	431	26.5	810	2	BG646074
34	429.2	26.3	817	2	BG583121
35	425.6	26.1	686	5	BQ860248
36	425.2	26.1	688	7	CV252809
37	425.2	26.1	691	7	CV252831
38	423.2	26.0	891	7	CK283034
39	420.8	25.8	435	1	AV794958
40	420.2	25.8	887	8	CK045487
41	418	25.7	698	2	BE433563
42	410.4	25.2	423	3	BP793989
43	409.8	25.2	765	7	CK274373
44	407.4	25.0	428	3	BP615264
45	406.8	25.0	421	3	BP606162

ALIGNMENTS

RESULT 1	CNS0A3HU	1537 bp	mRNA	linear	HTC 06-PBB-2004
CNS0A3HU	Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLSII442H09 of Silique of strain col-0 of Arabidopsis thaliana (thale cress).				
LOCUS	Arabidopsis thaliana				
DEFINITION	Arabidopsis thaliana (thale cress)				
ACCESSION	BX828861				
VERSION	BX828861.1	GI:42461179			
KEYWORDS	HTC; GSIT cDNA.				
SOURCE	Arabidopsis thaliana				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
AUTHORS	Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.				
TITLE	Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1537)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M. URGV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis Genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full-length http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.				
FEATURES	Location/Qualifiers				
SOURCE	1..1537				
	/organism="Arabidopsis thaliana"				
	/mol_type="mRNA"				

```
/db_xref="taxon:3702"
/clone="GSLT5IL442H09"
/tissue_type="Siliques"
/ecotype="Col-0"
/plasmid="pCMVSPORT 6"
complement(1..1537)
/gene="At4G34460"

ORIGIN

Query Match          92.0%; Score 1499.2; DB 4; Length 1537;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1512; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 13 CCGTGTGGTCTTGACATGATCTCTCTCAAGCTTTTAAATCTCTCTCTTTCCCA 72
Db 23 CCGTGTGGTCTTGACATGATCTCTCTCAAGCTTTTAAATCTCTCTCTTTCCCA 82
QY 73 CGTAATCCCCCAATCCATCTTTCTAGGGTTCGATCTCCCTCTCTCAATCATGACCT 132
Db 83 CGTAATCCCCCAATCCATCTTTCTAGGGTTCGATCTCCCTCTCTCAATCATGACCT 142
QY 133 TCTTCTCTTAGACCCCAAAAGTTTCCCTCTTTTATTGATCGCGACGAGAGCT 192
Db 143 TCTTCTCTTAGACCCCAAAAGTTTCCCTCTTTTATTGATCGCGACGAGAGCT 202
QY 193 AAGTCTGATCCCGGAATGTCTCTCCGAGCTCAAGAACGCCCGCTCGTACGGAG 252
Db 203 AAGTCTGATCCCGGAATGTCTCTCCGAGCTCAAGAACGCCCGCTCGTACGGAG 262
QY 253 ACCGTTAATACCTCCGTGACAGCTTAGACAGAGACGCCCTCAGCTCTCGATACCGAT 312
Db 263 ACCGTTAATACCTCCGTGACAGCTTAGACAGAGACGCCCTCAGCTCTCGATACCGAT 322
QY 313 GTGGCGAGGATTTACGGCGCGCAAGAGCTACTCGGGTGAGCTTCGGAGCAACGATCTG 372
Db 323 GTGGCGAGGATTTACGGCGCGCAAGAGCTACTCGGGTGAGCTTCGGAGCAACGATCTG 382
QY 373 GTTGTGTGCTGATCTTTACGGGACACACCGGAAAGTTTATTCATTAGATTGGACACCG 432
Db 383 GTTGTGTGCTGATCTTTACGGGACACACCGGAAAGTTTATTCATTAGATTGGACACCG 442
QY 433 GAGAGGAACCGGATTTGATGTCATCTCAAGATGGAGATTAATCGTGGGAATGCTCTA 492
Db 443 GAGAGGAACCGGATTTGATGTCATCTCAAGATGGAGATTAATCGTGGGAATGCTCTA 502
QY 493 ACGAGTCAGAAACTCATGCTATTAACCTCCCTTGTCATGGGTTATGACATGCTCTTC 552
Db 503 ACGAGTCAGAAACTCATGCTATTAACCTCCCTTGTCATGGGTTATGACATGCTCTTC 562
QY 553 TCTCCAAATGGTCACTCGGTTGCGTGTGGTATAGACAGTGTATGTTCTATCTTTAGC 612
Db 563 TCTCCAAATGGTCACTCGGTTGCGTGTGGTATAGACAGTGTATGTTCTATCTTTAGC 622
QY 613 CTTAGCTCAACGGCGGACAGGATCGAATGCTGACCGTTTCAAGATGCTACTGCTGAC 672
Db 623 CTTAGCTCAACGGCGGACAGGATCGAATGCTGACCGTTTCAAGATGCTACTGCTGAC 682
QY 673 AGGGATATGTTTCTGCTGTGATGTCCTCAAAATGAGGATGCCACCTATACCAAGT 732
Db 683 AGGGATATGTTTCTGCTGTGATGTCCTCAAAATGAGGATGCCACCTATACCAAGT 742
QY 733 TCAGTGTATCAAACTGTATCTTTATGGGATGPAATCTAGTGTCTCAAACTCTCTGTTTT 792
Db 743 TCAGTGTATCAAACTGTATCTTTATGGGATGPAATCTAGTGTCTCAAACTCTCTGTTTT 802
QY 793 GCGGTGATTTCACTCTGGACATCTGCTGATGATCAAGGCTCTCAATCAGTGGATCA 852
Db 803 GCGGTGATTTCACTCTGGACATCTGCTGATGATCAAGGCTCTCAATCAGTGGATCA 862
QY 853 AACCCAAACTGTTTATATCTGGTTATGCGGATTCACAGCACCGTTGTGGACACTCGT 912
Db 863 AACCCAAACTGTTTATATCTGGTTATGCGGATTCACAGCACCGTTGTGGACACTCGT 922
```

```
QY 913 GCTGCAAGCCGAGCGAGTCGCTACCTTTTCATCGTCAAGAGGAGATGTTTAATACGCTCAAG 972
Db 923 GCTGCAAGCCGAGCGAGTCGCTACCTTTTCATCGTCAAGAGGAGATGTTTAATACGCTCAAG 982
QY 973 TTTCTTCCGGATGGTATAGATTTGGGACTGGATCGACGATGGAACATGAGGCTGTAT 1032
Db 983 TTTCTTCCGGATGGTATAGATTTGGGACTGGATCGACGATGGAACATGAGGCTGTAT 1042
QY 1033 GACATAAGGACTGGTCAACCACTCCAGGTCATCAGCCACATGGTGTGAGAACCGA 1092
Db 1043 GACATAAGGACTGGTCAACCACTCCAGGTCATCAGCCACATGGTGTGAGAACCGA 1102
QY 1093 CTTGTCACTCCATTCGATCTCTGTGTACGGGAGACTTTCTTTTCGCTGGCTATCGGAGC 1152
Db 1103 CTTGTCACTCCATTCGATCTCTGTGTACGGGAGACTTTCTTTTCGCTGGCTATCGGAGC 1162
QY 1153 AACACACTTCTAGTTTGGGATACCTCTTGGGAGAGGTTGTATTCGATTTGGGATTA 1212
Db 1163 AACACACTTCTAGTTTGGGATACCTCTTGGGAGAGGTTGTATTCGATTTGGGATTA 1222
QY 1213 CAGCAGGATTCACACAGGAATAGAAATAGCTGTTTGGGTTGTCTCAGCAGATGGAAGTCA 1272
Db 1223 CAGCAGGATTCACACAGGAATAGAAATAGCTGTTTGGGTTGTCTCAGCAGATGGAAGTCA 1282
QY 1273 TTGTGTACAGGAATGGGATTTCAAAATCTAAAGATATGGCGTTTGGAGGACACAGGAGA 1332
Db 1283 TTGTGTACAGGAATGGGATTTCAAAATCTAAAGATATGGCGTTTGGAGGACACAGGAGA 1342
QY 1333 GTGATTTGAAGAAGATTTAACGAAAGTAGAGGTCACGCTCCAGTGTGTTGGTAAATATA 1392
Db 1343 GTGATTTGAAGAAGATTTAACG - AAAGTAGAGGTCACGCTCCAGTGTGTTGGTAAATATA 1401
QY 1393 TTTCTAGTCCGGAGTAAGGTTCCGTTTGTGGAAGGTTTGGTTTGAATAGTGGAGT 1452
Db 1402 TTTCTAGTCCGGAGTAAGGTTCCGTTTGTGGAAGGTTTGGTTTGAATAGTGGAGT 1461
QY 1453 GGTTAGAAGAAATTAACCTTCCCTTTTGTAGTGTGCTTTGATTTATTTATTTCTTCATTG 1512
Db 1462 GGTTAGAAGAAATTAACCTTCCCAATTTGTAGTGTGCTTTGATTTATTTATTTCTTCATTG 1521
QY 1513 GGAACTAAACTCCCTTC 1528
Db 1522 GGAACTAAACTCCCTTC 1537

RESULT 2
CNS0A3Q0
LOCUS
DEFINITION
CNS0A3Q0 1527 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTPG94ZC09 of Hormone Treated Callus of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION
BX828483
VERSION
BX828483.1 GI:42461795
KEYWORDS
HTC; GSLT_cDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 1527)
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
REFERENCE
2 (bases 1 to 1527)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
```


Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 1765)
Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.

Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

Unpublished
2 (bases 1 to 1765)
Genoscope.

Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

Location/Qualifiers

1..1765
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTSL382A01"
/tissue_type="Adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT.6"
complement(1..1765)
/gene="At4g34460"

ORIGIN

Query Match 76.8%; Score 1251.4; DB 4; Length 1765;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 1512; Conservative 0; Mismatches 1; Indels 240; Gaps 2;

22 GCTTTGACTGATCTCTCTCAAGCTTTTAAATCTCTCTCTCTTTTCCACGTAATTC 81

1 GCTTTGACTGATCTCTCTCAAGCTTTTAAATCTCTCTCTCTTTTCCACGTAATTC 60

82 CCCAAATCCATTTCTTAGGGTTCGATCTCCCTCTCTCAATCATGAACCTTTCTCTCT 141

61 CCCAAATCCATTTCTTAGGGTTCGATCTCCCTCTCTCTCAATCATGAACCTTTCTCT 120

142 CTAGACCCCAAGATTTCCCTTTTATTTGATCGGCGACGGAGAACCTTAAGTCTGAT 201

121 CTAGACCCCAAGATTTCCCTTTTATTTGATCGGCGACGGAGAACCTTAAGTCTGAT 180

202 CCCGGAATGTCGTCTCCGAGTCAAGAAGCCACGCCGCTCGCTACCGAGACCGTTAAT 261

181 CCCGGAATGTCGTCTCCGAGTCAAGAAGCCACGCCGCTCGCTACCGAGACCGTTAAT 240

262 AACCTCCGTGACCGAGCTTAGACAGAGACGCTCCAGCTCTCGATACCGAGTGTGCGAGG 321

241 AACCTCCGTGACCGAGCTTAGACAGAGACGCTCCAGCTCTCGATACCGAGTGTGCGAGG 300

322 TATTACGGCGCAAGGACGTACTCGGGTGAGCTTTCGGAGCAACCGATCTGGTTTGTGT 381

301 TATTACGGCGCGAAGGACGTACTCGGGTGAGCTTTCGGAGCAACCGATCTGGTTTGTGT 360

382 CGTACTCTTCAGGGACACACCGGA-----

1

Db 361 CGTACTCTTCAGGACACACCGGAAAGGTGACTCAAAAGTTTGGCATCTTTATCTCTTCAAGC 420

Qy 407 ----- 406

Db 421 ATGCTTCTTTTGTGGAGTGTTTTCTAGCTGTGTATATAAAAGCTTCCATTTTGGGGTGTAT 480

Qy 407 ----- 406

Db 481 CTAGTTTGTATTATTTAGTTGGTGTGTTTATAAGATGTTTACCAGTTATCATGTATGAGGAT 540

Qy 407 ----- 406

Db 541 TTGCTTAAGATCCTACCGGAATGTAGAGACAGGAGAGTTGATTGTGAATGTACAGTTGTT 600

Qy 407 -----AGGTTTATTTCATTAGATTGGACACCGGAGAGGAACC 442

Db 601 GTTGATTTAACATGTGGGATTTTGACAGTTTATTTCATTAGATTGGACACCGGAGAGGAACC 660

Qy 443 GGATTGTGTCAGTGCATCTCAAGATGGGAGATTAATCGTGTGGAATGCTCTAAACGATCAGA 502

Db 661 GGATTGTGTCAGTGCATCTCAAGATGGGAGATTAATCGTGTGGAATGCTCTAAACGATCAGA 720

Qy 503 AAACCTCATGCTATTAAACTCCCTTGTGCATGGGTTATGACATGCTTCTCTCCAAATG 562

Db 721 AAACCTCATGCTATTAAACTCCCTTGTGCATGGGTTATGACATGCTTCTCTCCAAATG 780

Qy 563 GTCAGTCGGTTGCGTGTGGATTTAGACAGTGTATGTTCTATCTTTAGCCCTTAGCTCAA 622

Db 781 GTCAGTCGGTTGCGTGTGGATTTAGACAGTGTATGTTCTATCTTTAGCCCTTAGCTCAA 840

Qy 623 CGGCGGACAGGATGGAACTGTACCGGTTTCAAGAAATGCTCACTGGTTCAGAGGGATATG 682

Db 841 CGGCGGACAGGATGGAACTGTACCGGTTTCAAGAAATGCTCACTGGTTCAGAGGGATATG 900

Qy 683 TTTTCGTGCTGTACGATGTCCCAATGAGGATGCCACCTTATCACCAGTTTCAGGTGATC 742

Db 901 TTTTCGTGCTGTACGATGTCCCAATGAGGATGCCACCTTATCACCAGTTTCAGGTGATC 960

Qy 743 AAACCTTGATCTTATGGGATGTAACTACTGTGCTCTCAAAAATCTTCTGTTTTTGGCGGTGAAT 802

Db 961 AAACCTTGATCTTATGGGATGTAACTACTGTGCTCTCAAAAATCTTCTGTTTTTGGCGGTGAAT 1020

Qy 803 TTGAGTCTGGACATATGCTGATGTACTAAGCGTCTCAATCAGTGGATCAAAACCAAACT 862

Db 1021 TTTAGTCTGGACATATGCTGATGTACTAAGCGTCTCAATCAGTGGATCAAAACCAAACT 1080

Qy 863 GGTTTATATCTGTTTCAATGGATTTCCACAGCACGGTTGTGGGACACTCGTCTGCAAGCC 922

Db 1081 GGTTTATATCTGTTTCAATGGATTTCCACAGCACGGTTGTGGGACACTCGTCTGCAAGCC 1140

Qy 923 GAGCAGTGCATACCTTTTCATGGTCAAGGAGAGATGTTAATACGGTCAAGTTCTTTCCGG 982

Db 1141 GAGCAGTGCATACCTTTTCATGGTCAAGGAGAGATGTTAATACGGTCAAGTTCTTTCCGG 1200

Qy 983 ATGGGTATAGATTTGGGATCGGATCAGATGGAACATGACAGCTGTATGACATAAGGA 1042

Db 1201 ATGGGTATAGATTTGGGATCGGATCAGATGGAACATGACAGCTGTATGACATAAGGA 1260

Qy 1043 CTGCTCACCACTCCAGGCTCTATCAGGCCATATGGTGTGGTGAACGGACCTGTCACT 1102

Db 1261 CTGCTCACCACTCCAGGCTCTATCAGGCCATATGGTGTGGTGAACGGACCTGTCACT 1320

Qy 1103 CCATTGCAATCTCTGTGTTCAGGGAGACTTCTTTTCGCTGGCTATGCGAGCAACCAACTT 1162

Db 1321 CCATTGCAATCTCTGTGTTCAGGGAGACTTCTTTTCGCTGGCTATGCGAGCAACCAACTT 1380

Qy 1163 GCTACGTTTGGGATACCGTCTTTGGGAGAGGTTGTATTTGGATTTGGGATTAACAGCAGATT 1222

Db 1381 GCTACGTTTGGGATACCGTCTTTGGGAGAGGTTGTATTTGGATTTGGGATTAACAGCAGATT 1440

Qy 1223 CACACAGGAATAGAAATAAGCTGTTTGGGTTGTGTCAGCAGATGGAAGTGCAATGTTGTACAG 1282

Db 1441 CACACAGGAATAGAAATAAGCTGTTTGGGTTGTGTCAGCAGATGGAAGTGCAATGTTGTACAG 1500

Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: meeki@rtc.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.	
FEATURES	Location/Qualifiers
source	1..663
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/clone="RAPL05-19-B08"
	/dev_stage="rosette plants"
	/lab_host="SOLR"
	/clone_lib="RAPLS"
	/note="Site 1: SstI; Site 2: XhoI; subjected to dehydration-treated(1,2,5,10,24 hr)"
ORIGIN	
Query Match 34.7%; Score 565; DB 1; Length 663; Best Local Similarity 98.2%; Pred. NO. 1.8e-155; Matches 601; Conservative 0; Mismatches 8; Indels 3; Gaps 3;	
QY	13 CGTGTGTTGTCCTGACGATCTCTCTCAAGCTTTTTTAAATCTCTCTCTTTTCCCA 72
DB	46 CGTGTGTTGTCCTGACGATCTCTCTCAAGCTTTTTTAAATCTCTCTCTTTCCCA 105
QY	73 CGTAATCCCCCAATCCATTTCTTAGGTTGATCTCCCTCTCTCAATCATGAACCT 132
DB	106 CGTAATCCCCCAATCCATTTCTTAGGTTGATCTCCCTCTCTCAATCATGAACCT 165
QY	133 TCCTCTCTTAGACCCACAAAGTTTCCCGCTTTTATTTGATCGGACGAGAGCCT 192
DB	166 TCCTCTCTTAGACCCACAAAGTTTCCCGCTTTTATTTGATCGGACGAGAGCCT 225
QY	193 AAGTCTGATCCCGAATCTGTCTCCGAGTCTCAAGAACCCAGCCGTCGCTACGGAG 252
DB	226 AAGTCTGATCCCGAATCTGTCTCCGAGTCTCAAGAACCCAGCCGTCGCTACGGAG 285
QY	253 ACCGTTAATAACCTCCGTGACAGCTTAGACAGAGACGCTCCAGCTCTCGATA-CGGA 311
DB	286 ACCGTTAATAACCTCCGTGACAGCTTAGACAGAGACGCTCCAGCTCTCGATACCGGA 345
QY	312 TGTGGCGAGGATTCAGCGGCGCAAGGACGTACTCGGTTGAGCTTCGGAGCAACGGATCT 371
DB	346 TGTGGCGAGGATTCAGCGGCGCAAGGACGTACTCGGTTGAGCTTCGGAGCAACGGATCT 405
QY	372 -GGTTTGTGTGCTACTCTTCAGGGACACCCGAAA-GGTTTATTCAATTAGTTGGACA 429
DB	406 GGGTGTGTGCTACTCTTCANGGACACCGAAAGGGTTTATTCAATTAGTTGGACA 465
QY	430 CCGGAGAGGAACCGGATGTTCAGTGCATCTCAAGATGGGAGATTAATCGTGTGGAATGCT 489
DB	466 CCGGAGAGGACCGGATGTGNCAGTGCATCTCAAGATGGGAGATTAATCGNGTGAATGCT 525
QY	490 CTAAGGATCAGAAACATCATGCTATTAAATCCCTTGTGCATGGGTTATGACATGTGCT 549
DB	526 CTAAGGATCAGAAACATCATGCTATTAAATCCCTTGTGCATGGGTTATGACATGTGCT 585
QY	550 TTCTCTCCAAATGGTCAGTGGTTCGCTGCTGGATAGACAGTGTATGTTCTATCTTT 609
DB	586 TTCTCTCCAAATGGTCAGTGGTTCGCTGCTGGATAGACAGTGTATGTTCTATCTTT 645
QY	610 AGCCTTAGCTCA 621
DB	646 AGCCTTAGCTAA 657
RESULT 6	
AV784390/c	
LOCUS	
DEFINITION AV784390 RAPL5 Arabidopsis thaliana cDNA clone RAPL05-19-B08 3', mRNA sequence.	
ACCESSION AV784390	
VERSION AV784390.1 GI:19803180	
KEYWORDS EST.	
SOURCE Arabidopsis thaliana (thale cress)	
ORGANISM Arabidopsis thaliana	
REFERENCE	
AUTHORS Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.	
1 (bases 1 to 590)	
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Ono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.	
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)	
JOURNAL Unpublished (2002)	
COMMENT Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: meeki@rtc.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a modified pBluescript vector as a SstI/XhoI insert. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.	
FEATURES	Location/Qualifiers
source	1..590
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"
	/db_xref="taxon:3702"
	/clone="RAPL05-19-B08"
	/dev_stage="rosette plants"
	/lab_host="SOLR"
	/clone_lib="RAPL5"
	/note="Site 1: SstI; Site 2: XhoI; subjected to dehydration-treated(1,2,5,10,24 hr)"
ORIGIN	
Query Match 34.0%; Score 554.4; DB 1; Length 590; Best Local Similarity 99.3%; Pred. No. 2.3e-152; Matches 588; Conservative 0; Mismatches 1; Indels 3; Gaps 3;	
QY	983 ATGGGTATAGATTTGGGACTGGATCAGACGATGGAAATCGAGGCTGTATGACATAAGGA 1042
DB	590 ATGGGTATAGATTTGGGACTGGATCAGACGATGGAAATCGAGGCTGTATGACATAAGGA 531
QY	1043 CTGGTCCCAACTCCAGGCTCTATCAGCCACATGCTGATGGTGAGACGGACCTGTCACT 1102
DB	530 CTGGTCCCAACTCCAGGCTCTATCAGCCACATGCTGATGGTGAGACGGACCTGTCACT 471
QY	1103 CCATTGCATTCTCTGTCTCAGGGAGACTCTTTTCGCTGGCTATGCCAGCAACAACACTT 1162
DB	470 CCATTGCATTCTCTGTCTCAGGGAGACTCTTTTCGCTGGCTATGCCAGCAACAACACTT 411
QY	1163 GCTACGTTTGGGATPACCTCTTGGGAGAGGTTGTATTGGATTTGGGATTTACAGAGGATT 1222
DB	410 GCTACGTTTGGGATPACCTCTTGGGAGAGGTTGTATTGGATTTGGGATTTACAGAGGATT 351
QY	1223 CACACGAGATAGATTAAGCTGTTTGGGTTGTTCAGCAGATGGAAATGTCATTGTGTACAG 1282
DB	350 CACACGAGATAGATTAAGCTGTTTGGGTTGTTCAGCAGATGGAAATGTCATTGTGTACAG 291
QY	1283 GAAGTGGGATTCAAATCTAAAGATATGGCGTTTGGAGGACACAGAGAGTGAATTTGAA 1342
DB	290 GAAGTGGGATTCAAATCTAAAGATATGGCGTTTGGAGGAC-CAGGAGATGAATTTGAA 232

QY 1343 GAAGATTTAACGAAAGTAGAGTACAGTCTCCAGTGTGTTGGTTAATATATATCTGTAGTC 1402
|||||
Db 231 GAAGATTTAACG-AAAGTAGAGTACAGTCTCCAGTGTGTTGGTTAATATATATCTGTAGTC 173
|||||
QY 1403 GCGAAGTAGGTTCCGTTTGGTGAAGTGTGTTGGTTTCAAAATAGTGGAGTGTAGAGA 1462
|||||
Db 172 GCGAAGTAGGTTCCGTTTGGTGAAGTGTGTTGGTTTGAATAGTGGAGTGTAGAGA 113
|||||
QY 1463 ATTAACCTCCCTTTTGTGA-GTGTGCTTTGATTTATTTATTTCTTCATTTGGAACTAAA 1521
|||||
Db 112 ATTAACCTCCCTTTTGTAGGTGTCTTTGATTTATTTATTTCTTCATTTGGAACTAAA 53
|||||
QY 1522 CTCCTTCAACGCTACTCAATGTAATCTCTGTAATCAATTTGTGTACCACC 1573
|||||
Db 52 CTCCTTCAACGCTACTCAATGTGAATCTCTGTAATCAATTTGTGTACCACC 1

RESULT 7
AY109568
LOCUS Zea mays CL2390_1 mRNA linear HTC 25-FEB-2005
DEFINITION Zea mays 1685 bp mRNA sequence.
ACCESSION AY109568
VERSION AY109568.1 GI:21213336
KEYWORDS HTC.
SOURCE Zea mays
ORGANISM Zea mays

REFERENCE
AUTHORS Gardiner, J., Schroeder, S., Polacco, M.L., Sanchez-Villeda, H., Fang, Z., Morgante, M., Landewe, T., Fongler, K., Uebeche, F., Hanafey, M., Tingey, S., Chou, H., Wing, R., Soderlund, C. and Coe, E.H. Jr.
TITLE Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization
JOURNAL Plant Physiol. 134 (4), 1317-1326 (2004)
PUBMED 15020742
REFERENCE 2 (bases 1 to 1685)
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.
TITLE Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes
JOURNAL Unpublished (2002)
REFERENCE 3 (bases 1 to 1685)
AUTHORS Coe, E.H.
TITLE Direct Submission
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at WSL, maizenap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES
Source Location/Qualifiers
1..1685
/organism="Zea mays"
/mol_type="mRNA"
/db_xref="MaizeDB:630952"
/db_xref="taxon:4577"
/clone_lib="Maize Mapping Project/DuPont Cornsensus Library"
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN

Query Match 33.9%; Score 552.2; DB 4; Length 1685;
Best Local Similarity 66.8%; Pred. No. 1.5e-151;
Matches 805; Conservative 0; Mismatches 388; Indels 12; Gaps 2;

QY 195 GTCTGATCCCGGAATGTCTGTCTCCGAGCTCAAAGAACGCCACGCCGCTCGCTACCGAGAC 254
|||||
Db 186 GCGCGACCCATGCGTCCGTGGCGGAGCTCAAGAGAAAGCACGCCGCGAGCTACGGCGTC 245
|||||
QY 255 CGTTAATAACCTCCGTGACAGCTTAGACAGAGAGCCCTCAGCTCCCTCCGATACCGATGT 314
|||||
Db 246 GGTGAACCTCCCTGCGGAGCGCTCCGCCAGCGCGGAGAGCGTCTCCGACACCGAGCGT 305
|||||
QY 315 GCGAGAGTATTTACGCGCGCAAGGACGTACTCGGTGAGCTTCGGAGCAACGGATCTGT 374
|||||
Db 306 GCGAGAGTACTCCAAGTCCGAGGGAGGGTCCGGTGAGCTTCAACCTACGGATCTGT 365
|||||
QY 375 TTGTTGCTGTACTCTTCAGGGACACACCGGAAAGGTTTATTTCAATTTAGATTGGACACCGA 434
|||||
Db 366 CTGCTGCCGCGACGCTGCAGGGGCCATAGCGGAAAGGTAATTTCTCTGGATTGGACCCCTGA 425
|||||
QY 435 GAGGAACCGGATTTGTGAGTGCATCTCAAGATGGGAGATTAATCGTGTGGAATGCTCTAAC 494
|||||
Db 426 AAGAAATTTGGATAGTCACTGCTCTCAAGATGGGAGGTTAATTTGTGGAAATGCAATTGAC 485
|||||
QY 495 GAGTCAGAAAACTCATGCTATTAACTCCCTTGTGCATGGGTTATGACATGTGCTTTCTC 554
|||||
Db 486 AAGCCAGAAACACATGCATAAGCTCATTTGCCATGGGTTATGGCGTGTGCTTTTTCG 545
|||||
QY 555 TCCAAATGCTCAGTCCGTTGGTGTGGTATAGACAGTGTATGTTCTTATCTTTAGCCT 614
|||||
Db 546 ACCCAATGCCAATCTGCTGCTGTGGTGTCTTGATAGTGGTCTCTATTTTCAATCT 605
|||||
QY 615 TAGCTCAACGGCGGCAAGGATGGAACTGTACCGGTTTCAAGAAATGCTCAGTGTGACAG 674
|||||
Db 606 CAATTTCTCAAGCAGACAGAGATGGCAACATGCCAGTATCAAGAATTTCTTACTGGACAAA 665
|||||
QY 675 GGGATATCTTTGCTGCTCTCAGTATGTCCTCAAAATCAGAGATGCCACCTTATCACCAGTTC 734
|||||
Db 666 GGGCTATGTCTCATCATGTCTCAATATGTCCAGATCAGGAAACACGCTTTATTACTAGTTC 725
|||||
QY 735 AGGTGATCAAACTGTATCTTATGGGATGTAATCTGCTCTCAAAATCTTCTGTTTTGG 794
|||||
Db 726 AGGTGATCAAACTGTATCTTGGGATGTTACTACTGGACAGAGATATCAATATTTGG 785
|||||
QY 795 CGGTGAATTTCAGTCTGACATATCTGCTGATGTAATAAGCGTCTCAATCAGTGGATCAAA 854
|||||
Db 786 TGGTGAATTTCCATCAGGGCATAACAGTCTGTTCAAAAGTGTGCTCATCAACTCAATCAAA 845
|||||
QY 855 CCCAACTGGTTTATATCTGTTTCAATGCGATTCACAGCACGGTTGTGGGACACTCGTGC 914
|||||
Db 846 TACAATATGTTGTCTCTGGCTCAATGATACAATCTGTGAGCTGTGGGATATCAGAA 905
|||||
QY 915 TCAAGCCGAGCAGTGCCTTACCTTTTCATGCTCAGCGGAGAGATGTTAATACCGTCAAGTT 974
|||||
Db 906 TCAAGTCAAGTGTTCGAACTTACCACTGAGGATGATGTTAAGTGTGAAGNN 965
|||||
QY 975 CTTTCCGATGGGTATAGATTTGGGACTGATCAGACAGATGGAACATGAGCGCTGTATGA 1034
|||||
Db 966 NNNCCCTGATGGCCATAGGTTTGGTACTGGCTCAGATGATGGCACAATGATATTTGA 1025
|||||
QY 1035 CATAGGACTGCTCACCACCTCAGGTCTATCAGCCACATGCTGATGG-----TGA 1085
|||||
Db 1026 TATGAGACAGGGCATCAACTTCAGGTGTACAGTAGGGAGCTGTAGAGAAATAGTAATGA 1085
|||||
QY 1086 GAACGACCTGTCACTCCATTCATTTCTCTGTGTCAGGAGACTTCTTTTTCGCTGGCTA 1145
|||||
Db 1086 ACTACTACTGTTATCATCTATTGCAATTTCAATATCAGGAAGCTACNNNNGCTGTTA 1145
|||||
QY 1146 TCGAGACAAACAACACTGTCTGATGTTGGGATACCTCTTTGGAGAGGTTGTATGGAATTT 1205
|||||
Db 1146 CTC---CAATGCTGACTGTTATGTGTGGGACACACTTCTCGCGAGGCTGTACTTAATTT 1202
|||||
QY 1206 GGGATTACAGCAGGATTTCACAGGAAATAGATAGCTGTTTGGGTTTCTCAGCAGATGG 1265
|||||

Qy	205	GGAATGCTGTCTCCGAGCTCAAAAGAACGCCAGCGCTCGCTACGGAGACCGTTAATAAC	264
Db	81	GAATGTGTCAGTTGGGAGCTGAAGAGCGGCACATGGCCGCTACACAGACTGTGAATGAT	140
Qy	265	CTCCGTGACCAGCTTTAGACAGAGACGGCTCCAGCTCTCTCGATACCGATGTGGCGAGGTAT	324
Db	141	CTCCGTGA AAAAATTAAAGAGAGCGCTCTCAATTACTCGACACAGATGTTCTGGGTAT	200
Qy	325	TCAGCGCGCAAGGACGTACTCGGGTGAGCTTCGGAGCAACGGATCTGGTTTCTTCTCGT	384
Db	201	GCAAGACGCAAGGTAAAAATCCGGTAACGTTTCGGCCCAACAGATCTAGTTTCTTCTAGG	260
Qy	385	ACTCTTCAGGACACACCGGAAAGTTTATTCAATTAGATTGGACACGGAGAGAAACCG	444
Db	261	ATCTCGAAGACACACTGGAAAGTCTATTCACTGGACTGGACTCTGAAAAAATCGT	320
Qy	445	ATTGTCAGTGCATCTCAAGATGGGAGATTAAATCGTGTGGAATGCTCTAACGAGTCAGAA	504
Db	321	ATAGTCAGTGCATCCCAAGATGGTAGATTAAATAGTGTGGAATGCTCTCAACGCCAGAA	380
Qy	505	ACTCATGCTATTAACCTCCCTTGTGCATGGGTTATGACATGTGCTTCTCTCCAAATGGT	564
Db	381	ACCCATGCAATTAAGCTTCCATGTGCTGGGTTATGACCTGTGCTCTCTCTCTAGTGA	440
Qy	565	CAGTCGGTTGGCTGTGCTGGATTAGACAGTGTATGTTCTATCTTTTAGCCCTTAGCTCAACG	624
Db	441	CAGTCTGTTGCTGTGGCGGCTTGA CAGTGCCTGCTCTATCTTCAACTTAATTCACCA	500
Qy	625	CGGACAAGGATGAACTGTACACGGTTTCAAGAAATGCTCACTGGTCA CAGGGGATATGTT	684
Db	501	ATCGATAAGGATGGATCCATCCAGTATCGAGAAATGCTTAGTGGGCATAAGGGGTATGTG	560
Qy	685	TCGTGCTGCTAGTATGCCAAATGAGSAGTCCCACTTATCACAGTTCAGGTGATCAA	744
Db	561	TC TTCGTGCTAGTATGTTCCGGATGAGGATATCACCCTTAATACCTAGTTCTGGTGATCAA	620
Qy	745	ACTTGTATCTTATGGGATGTAAC TACTGGTCTCAAAA CTTCGTTTGTGGCGTGAAATTT	804
Db	621	ACATGTGTACTTTGGGATATAACTACTTGGCCTAAGAACTTCTGTGTTTGGAGGTGAGTTT	680
Qy	805	CAGTCTGGACATAC TCTGTGATGTAACGGTCTCAATCAGTGGATCAAAACCACAATCGG	864
Db	681	CAATCTGGGCACATG CAGATGTATCAAGTGTCTCAATTAGTTCATCTAACCCCAAACTA	740
Qy	865	TTTATATCTGCTTCATGCGATTC CACAGCAGCGTTGTGGGACACTCGTGCTGCAAGCCGA	924
Db	741	TTTGTGTC TGGGTCTGTGACACAAC TCTCGACTGTGGGACACCCGAGTTGCTAGTCGA	800
Qy	925	GCAGTGCCTACTTTTCA TGGTCA CAGGGGAGATGTTAATACGGTCAAGTCTTTTCCGGAT	984
Db	801	GCTCAACGAA CATTTCATGSGGACAGGAGTGTATCTACTGTGTAAAGTTCCTTCCCTGAC	860
Qy	985	GGGTATACATTTGGGACTGGATCAGACGATGGAACATGCGGCTGTATGACATAAGGACT	1044
Db	861	GGTAATAGTTTGGAACTGGTT CAGATGATGCGAGCTGCAGATTATTTTGA CATTTAGACT	920
Qy	1045	GGTCA CCAACTCCAGGT --- CTATCAGCCA CATGGTGTGTGTGAGAACGGACCTGTCA	1099
Db	921	GGACAC CAGCTGCAAGTATACAA CCAACCGCATGGTGACGGTGCATVCCCTCATGTGA	978

RESULT 10	CO122720	858 bp	mRNA	linear	EST 16-JUN-2004
LOCUS	CO122720				
DEFINITION	GR_Eb04H21.f GR_Eb Gossypium raimondii cDNA clone GR_Eb04H21.5', mRNA sequence.				
ACCESSION	CO122720				
VERSION	CO122720.1	GI:48821407			
KEYWORDS	EST.				
SOURCE	Gossypium raimondii				
ORGANISM	Gossypium raimondii				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;				

```

rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium
1 (bases 1 to 958)
Kim.H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 04 row: H column: 21.
Location/Qualifiers
1. 858
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Eb04H21"
/tissue_type="floral"
/dev_stage="_3 to +3 DPA"
/lab_host="DH10B"
/clone_lib="GR_Eb"
/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; library made by Invitrogen with RNA supplied by
wented lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80

```

ORIGIN

Query Match	31.7%;	Score 516.2;	DB 7;	Length 858;
Best Local Similarity	76.8%;	Pred. No. 5.6e-141;		
Matches 658;	Conservative 0;	Mismatches 193;	Indels 6;	Gaps 2
QY	349	GTGAGCTTCGGAGCAACGGGATCTGGTTGTGTCTACTCTTCCAGGACACACCGGGAAG	408	
DB	4	GTTACTTTTGGTCTCAGGATCTGGTTGTTCAGGACCTTGCAGAGGTTCACACTGGCAAG	63	
QY	409	GTTTATTCTATTAGATTGGACACCGGAGAGGNAACGGGATTTGCAGTGCATCTCAAGATGGG	468	
DB	64	GTTTACTCTCTGGATTGGACTCCGGAAAGGNAACGGGATTTGTAGCGCGTCTCAGGATGGA	124	
QY	469	AGATTAATCGTGTGGAAATGCTCTAACGAGTCAGAAAACTCATGCTATTAAAACTCCCTTGT	528	
DB	124	CGATTAAATGTGTGACACGCTCTAACTAGCCAGAAAACTCAGCCATAAAGCTGCTTGT	183	
QY	529	GCATGGGTTATGACATGTGCTTTCTCTCCAAATGGTCAGTCGGTTGCGTGTGGTGATTA	588	
DB	184	CGTGGGGTTATGAGGTGTGCTTTTCTCCCAACGGGTCAATCAGATTGCTGTGGTGTCTC	243	
QY	589	GACAGTGTATGTTCTATCTTTAGCCTTTAGCTCAACGGCGGACAAAGGATGGAACTGTACCG	648	
DB	244	GATAGTATGTGTTTCAATTTTCACTTGAATTTCTCGACTGACAGGATGGAAACCTACT	303	
QY	649	GTTTCAAGAATGCTCACTGGTCAACAGGGGATATGTTTCGTGCTGTCAGTATGTCCCAAT	708	
DB	304	GTATCAAAAACCCCTTAGTGGACATAAGGGTTTACGTCCTCGTGTGTCAGTATGTTCCAGAT	363	
QY	709	GAGATGCCACCTTTATCACCAGTTCAAGTGATCAAACTGCTATCTTATGGATGTAATCT	768	
DB	364	GAAGACATTCAATATTACACGTTCCGGTGATCAGACCTGTGTTCTATGGGATATTACT	423	
QY	769	ACTGGTCTCAAAACTTCTGTTTTTTGGCGGTGAAATTTCACTGGAACATCTGCTGATGTA	828	
DB	424	ACTGGCCTTAGAAACATCGTTTTTCGGGGGAGAAATTTCAATCTGGACATACCTGCTGATGTG	483	
QY	829	CTAAGCGTCTCAATCAGTGGGATCAAAACCAACTGGTTTTATCTGGTTTCATGCCATCCC	888	
DB	484	CTAAGTGTCTCAATTAATGGGATCGAACTCAAGAAATGTTTGTCTCTGGTTCATGTGATGCA	543	
QY	889	ACAGCACGGTTGTGGGACACTCGTGTCTGCAAGCCGAGCAGTGCGTACCTTTTCATCGTCAAC	948	

Db 544 ACTGCCGATTTGGGATACTCGTGTGCAAGTCGAGCAGTTCGTACATTTTCATGGTCAC 603
Qy 949 GAGGAGATGTTAAATACCGTCAAGTCTTTTCCGGATGGGTATAGATTTGGGACTGGATCA 1008
Db 604 GAGGCGATGTAACACACCGTGAAGTCTTTCCAGATGGCAATAGGTTTGGAACTGGATCG 663
Qy 1009 GACGATGAACATGACGAGCTGTATGACATAAGGACTGTGTCACAACTCAAGT---CTAT 1065
Db 664 GATGATGGAACCTGTAGGTTGTTTGACATTAAGAACTGGTCAACCAATTACAGTACTAT 723
Qy 1066 CAGCCACATGTTGATGTTGAGAACGACCTGCACCTCAATTGCATTTCTCTGTGTCAGG 1125
Db 724 CAGCAACATGGGATTAACGAGGTTCCATGATGACCTCCATGCACTTCTCAATTCTGGA 783
Qy 1126 AGACTTCTTTTGGCTGGCTATGCGAGCAACACACTTGTCTAGCTTTGGGATACCCCTCTTG 1185
Db 784 AGACTTCTTTTGGCGATACCTC---AAATGGAGATTGCTAGTATGGGACAGGTTATTG 840
Qy 1186 GGAGAGGTTGATTGGA 1202
Db 841 GAAAAGGTCGTTTGA 857

RESULT 11

LOCUS CO072054 871 bp mRNA linear EST 15-JUN-2004
DEFINITION GR_Ea31A06.r GR_Ea Gossypium raimondii cDNA clone GR_Ea31A06 3',
mRNA sequence.

ACCESSION CO072054 GI:48741535

VERSION

CO072054.1

KEYWORDS

EST.

SOURCE

Gossypium raimondii

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 871)

REFERENCE

AUTHORS

Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,

Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and

Wing,R.A.

Global assembly of Cotton ESTs

Unpublished (2004)

Contact: Rod A. Wing

Arizona Genomics Institute

The University of Arizona

Forbes Building Room 303, Tucson, AZ, 85721-0036, USA

Tel: 520 626 9595

Fax: 520 621 1259

Email: rwing@genome.arizona.edu

Plate: 31 row: A column: 06.

Location/Qualifiers

1. 871

/organism="Gossypium raimondii"

/mol_type="mRNA"

/db_xref="taxon:29730"

/clone="GR_Ea31A06"

/tissue_type="whole seedlings"

/dev_stage="first true leaves"

/clone_lib="GR_Ea"

/note="Vector: pCMV.SPORT-6.1; Site 1: NotI; Site 2:

EcoRV; library made by Invitrogen with RNA supplied by

Wendle lab. Directional cloned into NotI-EV. Colonies

plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 31.6%; Score 515.4; DB 7; Length 871;

Best Local Similarity 75.4%; Pred. No. 9.8e-141;

Matches 654; Conservative 0; Mismatches 211; Indels 2; Gaps 1;

Qy 235 CACGCGTCGCTACGAGACGCTTAATAACCTCCCTGACCGCTTAGACAGAGCGCTC 294

Db 5 CACGCGCGCGACGAGACAGTCTCANTATCTTAAGAACAGATTAAAGCAAGCGTCAA 64

Qy 295 CAGCTCCTCGATACCGATGTGCGAGGTAATTCAGCGCGCAAGACGTACTCGGTGAGC 354

Db 65 CAGTTGCTCGACACTGATGTTGCTACATATGCAAGATCACAAGGAAGAGTCCAGTTACT 124

Qy 355 TTCGAGAACCGATCTGGTTGTTGCTGTAATCTTTCAGGAGACACACCGGAAGGTTTAT 414

Db 125 TTTGGTCTACCGATCTGCTTTGTTGAGGACCTTTCGAAGGTCACTCTGGAAGGTTTAC 184

Qy 415 TCATTAGATTGACACCGGAGAGAACCGGATTCGTAGTCGATCTCAAGATGGGAGATTA 474

Db 185 TCCTTGGATTGGACTCCGGAAGGAACCGGATTTGTAGCGGTCTCAGATGGACGATTA 244

Qy 475 ATCGTGTGAATGCTCTAAACGAGTCAGAAAACCTCATGCTATTAAACTCCCTTTGTCATGG 534

Db 245 ATTGTGGAACGCTCTAACTAGCCAGAAAACCTCACGCCATAAAGCTGCTTTGTCGTGG 304

Qy 535 GTTATGACATGTGCTTTCTCTCCAAATGGTCAGTCGGTGTGCTGTGGATTTAGACAGT 594

Db 305 GTTATGACGTGTGCTTTTCTCCCAACGGGTCAATCAGTTGCTGTGGTGTGCTCGATAGT 364

Qy 595 GTATGTTCTATCTTTAGCTTTAGCTCAACGCGGCAAGGATGGAACCTGTACCGGTTTCA 654

Db 365 ATGTTTCCATTTTCAACTTGAATTTCCGACTGACAGGGATGGAACCTACCTGATATCA 424

Qy 655 AGAATGCTCACTGGTTCACAGGGGATATGTTTTCGTGCTGTCTAGTATGTCCCAAATGAGAT 714

Db 425 AAAACCCCTTAGTGACATAAAGGTTACGTCCTCGTGTGTCAGTATGTTCCAGATGAAGAC 484

Qy 715 GCCACCTTATCACAGTTTCAAGTTGATCAAACTTGTATCTTTATGGGATGTAACCTGCT 774

Db 485 ATTCACATTTATACAGTTTCGGTGCATCAGACCTGTGTTCTATGGGATATTTACTACTGGC 544

Qy 775 CTCAAACTCTCTGTTTGGCGGTGAATTTTCAGCTGCGACATCTGCTGATGTAAGTACTAGC 834

Db 545 CTTAAGAACGTTTTTCGGGGGAGAAATTTCAATCGACATATCTGCTGATGTTGCTAAGT 604

Qy 835 GTCTCAATCAGTGGATCAAAACCCAAACTGGTTTATATCTGTTTCATGCGATTCCACAGCA 894

Db 605 GTCTCAATTAATGATCGAACTCGAACTCAAGAATGTTTGTCTCTGTTTCATGTGATGCAACTGCC 664

Qy 895 CGTTTGGGACACTCGTGTGCAAGCGGACGACGTCGCTTTTCATGTCACGAGGA 954

Db 665 CGATTGTTGGGATCTCGTGTGCAAGTCGAGCAGTTCTGATACATTTTCATGTCACGAGGCG 724

Qy 955 GATGTTAATACGCTCAAGTTCTTTCCGATGGGTATAGATTGCGACTTGGATCAGACCAT 1014

Db 725 GATGTAACACCGTGAAGTTCTTTCCAGATGGCAATAGTTTGGAACTGGATCGGATGAT 784

Qy 1015 GGAACATGCGAGGCTGTATGACATAAGGACTGGTCACCAACTCCAG--GTCTATCAGCCAC 1072

Db 785 GGAACCTGTAGTTGTTTGACATTTAGAACTGGTCAATTAAGTATACATGATATCTATCAGCAAC 844

Qy 1073 ATGTTGATGTTGAGAACGGACCTGTCA 1099

Db 845 ATGCGGATAACGAGGTTCCACATGTGA 871

RESULT 12

AI994586

LOCUS

DEFINITION

701498643 A. thaliana, Ohio State clone set Arabidopsis thaliana

cDNA clone 701498643, mRNA sequence.

ACCESSION

AI994586

VERSION

AI994586.1

KEYWORDS

EST.

SOURCE

Arabidopsis thaliana (thale cress)

ORGANISM

Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 535)

REFERENCE

AUTHORS

Chen,J., Momiyama,M., Chan,E., Mooney,M., Carroon,B., Gilliland,D.,

Wang,X., Hillman,J., Guegler,K., Kim,C., Doyle,M., Brzaska,P., Gorgone,G., Burns,D., Griffin,J., Mouanoutoua,M., Nguyen,D., Tan,R., Rose,M., Warren,B., Ton,B., Kastury,K., Borillo,C., Carpio,T., Policky,J., Suzuki,G., Argentine,C., Shah,S., Nobriga,A., Murry,L., Turner,C., Krikorian,S., Elder,L. and Hanson,D.

Arabidopsis thaliana Gene Expression Microarray
Unpublished (1999)
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

Location/Qualifiers

1. 535
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone_lib="A. thaliana, Ohio State clone set"
/note="cDNA library was made from selected clones from the Arabidopsis thaliana Ohio State clone set."

ORIGIN

Query Match 31.5%; Score 513.2; DB 1; Length 535;
Best Local Similarity 97.6%; Pred. No. 3.7e-140;
Matches 521; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 71 CACGTAATCCCGCAATCCATCTTTCTAGGGTTCGATCTCCCTCTCATCATGATGAC 130
Db 1 CACGTAATCCCGCAATCCATCTTTCTAGGGTTCGATCTCCCTCTCATCATGATG 60
QY 131 CTCTCTCTCTCTAGACCCCAAAAGTTTCCTCCCTTTTATTGATCGGCGAGAGAGC 190
Db 61 CTCTCTCTCTCTAGACCCCAAAAGTTTCCTCCCTTTTATTGATCGGCGAGAGAGC 120
QY 191 CTAAGTCTGATCCCGAATGTCTGTCTCGAGCTCAAGAAGCCACCGCGCTGATACGG 250
Db 121 CTAAGTCTGATCCCGAATGTCTGTCTCGAGCTCAAGAAGCCACCGCGCTGATACGG 180
QY 251 AGACGGTTAATACCTCGTGACAGCTTAGACAGAGAGCCCTCAGCTCTCGATACCG 310
Db 181 AGACGGTTAATACCTCGTGACAGCTTAGACAGAGAGCCCTCAGCTCTCGATACCG 240
QY 311 ATGTGCGGAGGTATTACGCGCGCAAGGACGTACTCGGGTGAGCTTCGGAGCAACGGATC 370
Db 241 ATGTGCGGAGGTATTACGCGCGCAAGGACGTACTCGGGTGAGCTTCGGAGCAACGGATC 300
QY 371 TGGTTTGTGTCGTACTCTTCAGGGACACACCGGAAAGGTTTATTCAATTAGATTGGACAC 430
Db 301 TGGTTTGTGTCGTACTCTTCAGGGACACACCGGAAAGGTTTATTCAATTAGATTGGACAC 360
QY 431 CGGAGGAAACCGGATGTTCAGTGCACTCAAGATGGGAGATTAATCGTGGGATGCTC 490
Db 361 CGGAGGAAACCGGATGTTCAGTGCACTCAAGATGGGATTAATCGTGGGATGCTC 420
QY 491 TACGAGTCAGAAACCTCATGCTATTAACTCCCTTGTGATGGGTTATGACATGCTT 550
Db 421 TACGAGTCAGAAACCTCATGCTATTAACTCCCTTGTGATGGGTTATGACATGCTT 480
QY 551 TCTCTCCAAATGGTTCAGTGGTTCGTTGGTGGATTAGACAGTGTATGTTCTTA 604
Db 481 TCTCTCCAAATGGTTCAGTGGTTCGTTGGTGGATTAGACAGTGTATGTTCTTA 534

RESULT 13
CO072562
LOCUS GR_Ea31M19_r GR_Ea Gossypium raimondii cDNA clone GR_Ea31M19 3',
DEFINITION mRNA sequence.
ACCESSION CO072562

CO072562.1 GI:48742043
EST.
Gossypium raimondii
Gossypium raimondii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 838)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
Global assembly of Cotton ESTs
Unpublished (2004)
Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 31 row: M column: 19.
Location/Qualifiers
1. 838
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone_lib="GR_Ea31M19"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_lib="GR_Ea"
/note="Vector: PCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 31.4%; Score 511.4; DB 7; Length 838;
Best Local Similarity 76.2%; Pred. No. 1.5e-133;
Matches 629; Conservative 0; Mismatches 196; Indels 0; Gaps 0;

QY 235 CACGCGTTCGTACGAGACCGTTAATACCTCCGTCAGCTTAGACAGAGCGCTC 294
Db 5 CACGCGCGGCGACGAGACAGTGAATTAATCTAAGAGAACGTTAAGCAAGAGCTCAA 64
QY 295 CAGCTCTCTCGATACCGATGTGCGGAGGTATTACGCGCGCAAGAGACGTACTCGGCTGAGC 354
Db 65 CAGTTGCTCGACACTGATGTTGCTACATATGCAAGATCACAAGGAGAGTCCAGTTACT 124
QY 355 TTCGAGCAACCGATCTGGTTTGTGCTACTCTTCAGGAGACACACCGGAAAGGTTTAT 414
Db 125 TTTGCTCTACCGATCTGGTTTGTTCGAGGACCTTTCAGGAGTCACTGGCAAGGTTTAC 184
QY 415 TCATTAGATTGACACCGGAGGAAACCGATTGTTCAGTCACTCTCAAGATCGGAGATTA 474
Db 185 TCCTTGGATTGACCTCCGAAAGGAAACCGATTGTTCAGGCTCTCAGGATCGACGATTA 244
QY 475 ATCGTGTGGAATGCTCTAAACGAGTCAGAAAACCTCATGCTATTAAATCTCCCTTGTGATGC 534
Db 245 ATTGTGTGGAACGCTCTAACTAGCCAGAAAACCTACGCCATAAAGCTGCTTGTGCGTGG 304
QY 535 GTTATGACATGTCTTCTCTCCAAATGGTTCAGTGGTTCGTTGGTGGATTAGACAT 594
Db 305 GTTATGACATGTCTTCTCTCCAAATGGTTCAGTGGTTCGTTGGTGGATTAGACAT 364
QY 595 GTATGTTCTATCTTTAGCTTAGCTCAACGGCGGCAAGGATGGAACCTGACCGGTTTCA 654
Db 365 ATGTGTTCAATTTTCAACTTGAATTTCTCGACTGACAGGATGGAACCTGACCTGTATCA 424
QY 655 AGAATGCTCAGCTGGTCACAGGGGATATGTTTCGTGCTGTTCAGTATGTCCCAATAGGAT 714
Db 425 AAAACCCCTTAGTGGACATAAGGGTTACGCTCTCGTGTGTCAGTATGTTCCAGATGAAGAC 484

```
QY 715 GCCACCTTATCACAGTTTCAGGTGATCAAACTTGTAICTTTATGGGATGTAACACTACTGGT 774
Db |||
485 ATTCAATATTACAGTTTCGGGTGATCAGACCTGTGTTCATGGGATATTACTACTGCG 544
QY 775 CTCAAAATCTCTGTTTTTGGGGTGAATTTCACTGTCGACATACCTGCTGATGTAACG 834
Db |||
545 CTTAGAACATCGTTTTTCGGGGAGAAATTTCAATCTGGACATACCTGCTGATGTCGTAAGT 604
QY 835 GTCTCAATCAGTGGATCAAAACCAACTGGTTTATATCTGGTTCAATGCGATTCCACAGCA 894
Db |||
605 GTCTCAATTAATGGATCGAATCTCAAGAATGTTTGCTCTGTTTCATGTCGATGCAACTGCC 664
QY 895 CGGTTGTGGGACATCTGCTGCTGCAAGCCGACAGTGGCTACCTTTTCATGGTTCACAGGGA 954
Db |||
665 CGATTGTGGGATACCTCGTGTGTGCAAGTCGACAGTTCGTACATTTTCATGGTTCACAGGGC 724
QY 955 GATGTTAATACGTCGAAGTCTTTCCGGATGGTATAGATTGGGACTGGATCGACAGCAT 1014
Db |||
725 GATGTTAAACACCGTGAAGTCTTTCCAGATGGCAATAGTTTGGAACTGGATCGGATGAT 784
QY 1015 GGAACATCAGGCTGTATGACATAGAGACTGCTACCAACTCCAG 1059
Db |||
785 GGAACCTTAGGTTGTTGACATAGAACTGCTCAATTAAG 829

RESULT 14
AM058772 645 bp mRNA linear EST 11-AUG-2005
LOCUS AM058772 AAFc WHRI_BoE01a Brassica oleracea var. alboglabra cDNA
DEFINITION clone AAFc_WHRI_BoE01a009E15_SP6, mRNA sequence.
ACCESSION AM058772
VERSION AM058772.1 GI:72287665
KEYWORDS EST.
SOURCE Brassica oleracea var. alboglabra (Chinese kale)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 645)
AUTHORS Barker,G.C., Sharp,A., Durkin,J., Naylor,R., Edwards,R.,
Gjetvaj,B., Lydiate,D. and King,G.
TITLE Analysis of EST libraries of Brassica oleracea prepared from
differing stages of development
JOURNAL Unpublished (2005)
COMMENT Contact: Barker GC
WHRI
Warwick University
Wellesbourne, Warwick, CV35 9EF, UNITED KINGDOM.
FEATURES
source
Location/Qualifiers
1..645
/organism="Brassica oleracea var. alboglabra"
/mol_type="mRNA"
/cultivar="A12"
/db_xref="taxon:3714"
/clone="AAFc_WHRI_BoE01a009E15_SP6"
/tissue_type="etiolated seedling"
/clone_lib="AAFc_WHRI_BoE01a"

ORIGIN
Query Match 31.1%; Score 506.6; DB 1; Length 645;
Best Local Similarity 89.4%; Pred. No. 3.6e-138;
Matches 542; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 700 GTCCCAATGAGGATGCCACCTTATACCACTTCAGGTGATCAAACTGTATCTTATGG 759
Db |||
40 GTCCCAATGAGGATGCTGCTGTATCAAGTTGAGTATCAAACTGTGCTTATGG 99
QY 760 GATGTAATCTAGTGTCTCAAACTTCTGTTTTTGGCGGTGAATTTCACTCTGGACATACT 819
Db |||
100 GACGTAACTACTGGGCTCAAACTTCTGTTTTTGGTGGCGAGTTTCAGTCTGGACATACT 159
QY 820 GCTGATGTAAGGTTCTCAATCAGTGTGATCAAACTGGTTTATATCTGTTCA 879
Db |||
```

```
Db 160 GCTGATGTACTAAGTGTCTCAATCAGTGGATCAAAACCAAACTGTTTCATATCTGTTCA 219
QY 880 TGGCATTTCCACGACCGTTCGGGACACTCGTCTGCAAGCCGACGAGTGCCTACCTTT 939
Db |||
220 TGGACACACCGACCGCTTGTGGGACACTCGCGCTGGAGCCGAGCGGTGCGGACGTTT 279
QY 940 CATGGTCAAGGAGAGATGTTAAATACCGGTCAAGTTCCTTTCCGGATGGGTATAGATTTGGG 999
Db |||
280 CATGGCACAGGAGAGATGTTAAATCTGTCAAGTTCCTTTCCGGATGGGATAGATTTGGG 339
QY 1000 ACTGGATCAGACGATGGAACATGACGAGCTGTATGACATAGGACTGTCACCACTCCAG 1059
Db |||
340 ACTGGATCAGAAGATGGAACATGACGAGTTGTATGACATTAGGACCGGTTCATCAACTCCAG 399
QY 1060 GTCTATCAGCCACATGTTGATGGTGAGAACCGAGCTGTCACTTCATTTGCAATTTCTCTGTG 1119
Db |||
400 GTGTATCAGCCACATGTTGATGGTGAGAACGCTACTCTGTGCTCAATTTGCTGCTGCC 459
QY 1120 TCAGGAGACTTCTTTTCGCTGCTATGCGAGCAACCACTTCTGCTAGTTTGGGATACC 1179
Db |||
460 TCGGGAGACTTCTTTTCGCTGGATACGCTAAACAACGCTGCTAGTTTGGGATACG 519
QY 1180 CTCTTGGGAGAGTGTATTTGGATTTGGGATTTACAGCAGGATTTACACAGGAATAGATA 1239
Db |||
520 CTTTGGGAGAGATTTGATTTGGATTTGGGGAGCTGCAGGATTCGCACAAGATCGGATA 579
QY 1240 AGCTGTTTGGGTTGTCAGCAGATGGAAGTGCATTTGTACAGGAAGTTGGGATTTCAAT 1299
Db |||
580 AGCTGTTTGGGGTGTGACGNNNNGAAGTCCCTGTGTACAGGAAGCTGGGATTTCAAT 639
QY 1300 CTAAG 1305
Db |||
640 CTAAG 645

RESULT 15
DN620362 778 bp mRNA linear EST 22-MAR-2005
LOCUS UCRCS11_05D10 f Parent Washington Navel Orange Scale-Infested Rind
DEFINITION cDNA Library UCRCS11 Citrus sinensis cDNA clone CS_WED0005D10, mRNA
sequence.
ACCESSION DN620362
VERSION DN620362.1 GI:61689462
KEYWORDS EST.
SOURCE Citrus sinensis
ORGANISM Citrus sinensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Sapindales; Rutaceae; Citrus.
REFERENCE 1 (bases 1 to 778)
AUTHORS Close,T.J., Roose,M.L., Federici,C.F., Mandal,J., Fenton,R.D.,
Luck,R., Forster,L., Morse,J., Flores,P., Wanmaker,S., Kim,H.R.,
Kudrna,D., Collura,K., Wissotski,M., Byrne,M., Stum,D., Smart,D.,
Muller,C. and Wing,R.
TITLE Development of EST Resources and New Genetic Markers for California
Citrus - Parent Washington Navel Orange Scale-Infested Rind cDNA
Library UCRCS11
JOURNAL Unpublished (2005)
COMMENT Contact: Timothy J. Close
Department of Botany & Plant Sciences
University of California
Riverside, CA 92521-0124, USA
Tel: 909-787-3318
Fax: 909-787-4437
Email: timothy.close@ucr.edu
Seq primer: T7.
Location/Qualifiers
1..778
/organism="Citrus sinensis"
/mol_type="mRNA"
/cultivar="Parent Washington Navel"
/db_xref="taxon:2711"
/clone="CS_WED0005D10"
```


/tissue_type="Flavado, albedo, some red scale"
 /dev stages="12 year old trees"
 /lab host="E. coli TJC121"
 /clone lib="Parent Washington Navel Orange Scale-Infested
 Rind cDNA Library UCRCS11"
 /note="vector: Lambda Uni-ZAP XR, excised phagemid;
 Site_1: EcoRI; Site_2: XhoI; In cooperation with Dr.
 Robert Luck and Dr. Joseph Morse, Department of
 Entomology, University of California Riverside, navel
 orange fruit was infested with red scale (Aonidiella
 aurantii [Maskell]). Dr. Luck's technician, Lisa Forster,
 maintained colonies of red scale. Dr. Morse's technician,
 Paul Flores, cleaned and infested the fruit, bagged the
 fruit after the crawlers had time to settle, then
 harvested the fruit at the early second instar and late
 third instar phases of the insect. Claire Federici (Dr.
 Mikeal Roose's lab) cleaned off most of the insects (some
 remained firmly attached), cut the peel, then froze and
 stored at -80°C. The trees used for this set of samples
 were growing in field 16k of the University of California
 Citrus Experiment Station. This was the same field from
 which the tissue was collected for the navel shoot
 meristems and the peel for citrus thrips. Scitotothrips
 citri [Moulton]. The trees were planted 12 October 1992.
 The scion was Parent navel. The trees used for infesting
 with red scale all had Troyer rootstock. Three different
 trees were used for each of the two infestations, a total
 of 6 trees. This experiment did not take place according
 to the original schedule because the red scale colony
 became infested with mites, so too few insects were
 available to infest the fruit on the date originally
 planned. The plan had been to infest on or about June 30
 and August 13. Instead the infestation dates were August
 30 and September 27, 2004. Because this took us into the
 cool weather of fall, the insects applied on the second
 infestation date took about three times as long to develop
 to the same stage as the insects applied on the first
 infestation date. The insects were reared on lemons.
 Before infesting the fruit on the trees, Paul Flores
 cleaned the peel and checked to make sure it was not
 already infested. He placed about 200 crawler stage
 insects on the fruit using a soft paintbrush. After the
 crawlers had a day to move around and find a spot to
 attach, Paul placed a very fine mesh drawstring bag over
 the fruit to exclude predators and parasitoids of the
 scale. Bags were left in place for the duration of the
 field development period and also covered control fruit.
 Paul Flores infested fruit on 30 August and 1 September
 2004 after cleaning the fruit on August 27. On 21
 September the first set of red scale infested fruit were
 sampled. He brought Claire Federici 15 infested fruit and
 12 uninfested control fruit. The insects were at the early
 second instar. The controls had been cleaned and bagged at
 the same time as inoculated ones, but had no scale
 introduced. The infestation was heavy, so the insects were
 not well separated on the fruit. A razor blade was used to
 slice off the flavado with about half the thickness of the
 albedo included. The peel from only the stem half of the
 fruit was used because that was where the insects were
 most concentrated. The insects were impossible to wash off
 without severely damaging the peel; each fruit was washed
 with water and a sponge or bottlebrush and wiped dry with
 a paper towel before cutting it, but many of the insects
 were still intact and were included in the sample. The
 pulp had color but the fruit rind was still green. The
 peel from each half fruit was bagged separately in foil
 packets, and pressed between sheets of dry ice to freeze.
 All packets were placed in a paper bag with the date and
 information about 'control' or 'infested' written on the
 bag, and then placed at -80°C. The second sampling date
 from the first infestation was 11 October. Paul brought 19
 infested and 9 uninfested fruit to Claire. These were all
 set up on 30 August. The insects were at the late third

instar. Mikeal Roose's advice was to prepare three bulks
 of each, 5 infested and 3 uninfested fruit per bulk. The
 remaining four infested fruit were infested to a lesser
 degree and were discarded. Washing was done as before and
 the peel from each fruit frozen between sheets of aluminum
 foil pressed between sheets of dry ice. There were quite a
 lot of adhering insects in the frozen rind. After it was
 frozen, the peel from 3 uninfested or 5 infested fruit
 were pooled in a foil packet, then stored at -80°C. The
 pulp had color but the rind was mostly green. However,
 there was a little color developing on some fruit around
 the insects. The third sampling date was 2 December. Paul
 brought the first fruit from the second infestation, which
 were set up on September 27 or 28. He brought 15 infested
 and 10 uninfested fruit. Claire cut and froze them as
 before, making three pools of 5 infested fruit and three
 pools of 3 uninfested fruit. The control fruit all had
 green at the stem end, mostly about 1/4 of the
 circumference from stem to blossom end, but one fruit was
 green all over. The infested fruit were almost fully
 orange, with only a small amount of the surface still
 green, approximately the size of a quarter coin. Paul said
 it would be at least a month, probably 6 weeks or even
 more before the insects reach the late third instar, so
 this was the final sample for the cDNA library due to time
 constraints for EST sequencing. Mandal and Fenton (Close
 lab) purified RNA by the phenol method described in J.
 Japanese Soc. Hort. Sci. 1996. 64 (4): 809-814, purified
 poly(A) mRNA using PolyAT Tract mRNA Isolation Kit
 (Promega), produced a primary cDNA library using a lambda
 ZAP XR cDNA Synthesis Kit (Stratagene), then mass-excised
 1 million pfu from the primary library to produce a
 phagemid population. The library was made from a mixture
 of RNA from each of the three treatments such that
 approximately equal amounts of early second instar and
 late third instar RNA were used. Phagemids were plated,
 plasmid DNA purified, cDNA clones archived, and DNA
 sequences determined bi-directionally using an ABI3730 at
 the Arizona Genomics Institute (Kim, Kudrna, Collura,
 Wisotzki, Byrne, Stum, Muller, Wing). Chromatogram
 files were downloaded by FTP by Close, then processed by
 Wamamaker (Close lab) using the HarVest pipeline
 (http://harvest.ucr.edu) to remove vector and cloning
 oligo sequences and various contaminants, and to trim to a
 high quality region. Sequences that retained a phred 17
 region of at least 100 bases were assembled, then chimeras
 were removed following manual inspection of assemblies
 (Wamamaker, Close, Roose). Sequences that survived all
 removal steps were submitted to GenBank. Clones from this
 library are archived at the Arizona Genomics Institute
 (http://www.genome.arizona.edu/orders/).

ORIGIN

Query Match	31.0%;	Score	504.6;	DB	8;	Length	778;
Best Local Similarity	79.1%;	Pred.	No. 1.5e-13;				
Matches	613;	Conservative	0;	Mismatches	159;	Indels	3;
Gaps	1;						

QY	325	TCAGCGGGCGAAGGACGTA	CTCGGGTGAGCTTCGGAGCAACGGATCTGGTTTGTGCT	384
DB	4	TCGAGGTCGACGGGTGGGCT	CCGGTCAGCTTGGTCCGACCGATCTAGTTGTTGTAGG	63
QY	395	ACTCTTCAGGGACACAC	CGGAAAGGTTTATTCATTAGATTGGACACCGGAGGAGCCGG	444
DB	64	ACCTTGAAGGCCACGCT	GGCAAGGTATATTCATTGGACTGGACTTCCGAAAGGAAATCGA	123
QY	445	ATTGTGAGTGCATCT	CAAGATGGGAGATTAAATCGNGTGAATGCTCTAACAGAGTCAGAAA	504
DB	124	ATTGTGAGTGCATCT	CAAGATGGGAGATTAAATGAGTGTGAATGCTCTTACAAGCCAGAAA	183
QY	505	ACTCATGCTATTAAACT	CCCTTGTGTCATGGGTTATGACATGTGCTTTCTCTCCAATGCT	564
DB	184	ACTCATGCTATAAANA	CTGCTTGTGGTAAATGACTTGTGCTCTTCTCTCCAATGCT	243

QY	565	CAGTCGGTTGCGTGGTGGATTAGACAGTGATGTTCTATCTTTAGCCTTAGCTCAACG	624
Db	244	CAATCTGTGCTGTGGTGGTCTTGATAGTGTTCCTCCATTTTCAACCTGAATTCCTCT	303
QY	625	GCGGACAAAGGATGGAACCTGTACCGGTTTCAAGAACTCTACCTGGTCACAGGGGATATGTT	684
Db	304	ACTGACAAAGGACGGGAATTTACAGATATCAAGATGCTTAGTGGGCACAAGGGTTATGTG	363
QY	685	TCGTGCTCTCAGTATGTCCCAAAATAGGATGCCACCTTATCACCAGTTTCAGGTGATCAA	744
Db	364	TCCTGCTGTCTAGTATGTTCTTGATGAGGACACTCACCTAAATFACCAGTTTCAGGTGACCAA	423
QY	745	ACTTGTATCTTATGGGATGTAACACTACTGGTCTCAAAACTTCTGTTTTTGGCGGTGAATTT	804
Db	424	ACCTGTGTTTTTGGGATATTAATAAGGCTTAGAACTTCTGTTTTTGGAGGAGAAATTT	483
QY	805	CAGTCTGGACATACCTGCTGATGTACTAAGCGTCTCAATCAGTGGATCAAACTCAAACTGG	864
Db	484	CAGTCTGGGCACACTGCTGATGTGTTAAGTGTCTCAATAGTGGTCAAACTCAAGGATG	543
QY	865	TTTATATCTGGTTTCATGCGATTCCACAGCAGCGTTGTGGGACACTCGTGTGCAAGCCGA	924
Db	544	TTTGTGTCTGGTTCCTGTGATTCTACTGCCGATTGTGGGATCTCGTGTGCAAGTGA	603
QY	925	GCAGTGGCTACCTTTTCATGGTCAAGAGGAGATGTTAATACGGTCAAGTCTTTCGGAT	984
Db	604	GCAGTGGCCACATTTTCATGGGCATGAGGGAGATGTTAAACACTGTGAAGTCTTTCAGAT	663
QY	985	GGGTATAGATTTGGGACTGGATCAGACGATGGAACATGCAGGCTGTATCAGATAGGACT	1044
Db	664	GGCAATAGTTTGGAACTGGCTCAGACCATGGAACCTTGCAGATTTATTTGATATCAGGACT	723
QY	1045	GGTCACCAACTCCAGT---CTATCAGCCACATGGTGTATGGTGAGAACGGACCTG	1096
Db	724	GGGCACCAACTCCAAGTGTATTATCAGCAGCAGCGGTGAGAAATGAGATCCACATG	778

Search completed: December 28, 2005, 19:27:18
Job time : 5985 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 28, 2005, 14:53:47 ; Search time 319 Seconds
(without alignments)
9077.266 Million cell updates/sec

Title: US-10-602-898A-1
Perfect score: 1629
Sequence: 1 cctgacgtacgcgtgttg.....cagaagataaaacgtacta 1629

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/1_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/H_COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCWUS_COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE_COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfileseq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	130	8.0	1520	3	US-09-492-029-2
2	110.8	6.8	1523	3	US-09-803-653-1
3	110.8	6.8	1523	3	US-09-492-029-4
4	109.4	6.7	3078	3	US-09-949-016-1956
5	109.4	6.7	3088	3	US-09-023-655-1294
6	109.2	6.7	1517	3	US-09-180-783-1
7	109.2	6.7	1523	3	US-09-803-653-2
8	107.8	6.6	1542	3	US-09-949-016-1365
9	96	5.9	894	3	US-09-803-653-4
10	83.8	5.1	1394	3	US-09-147-826B-1
11	71.4	4.4	2119	3	US-09-032-372-6
12	61	3.7	659	3	US-09-016-434-1017
13	58.2	3.6	551	3	US-09-270-767-53
14	58.2	3.6	551	3	US-09-270-767-5335
15	52.8	3.2	110243	3	US-09-949-016-13698
16	51.8	3.2	1215	3	US-09-248-796A-6112
17	50.6	3.1	7218	2	US-08-232-463-14
18	44.8	2.8	7218	2	US-08-232-463-14
19	44.2	2.7	9308	3	US-09-949-016-13107
20	42.8	2.6	2359	2	US-08-188-582-4
21	42.8	2.6	2359	2	US-08-646-715-4
22	42.4	2.5	1422	3	US-09-248-796A-4567
23	41	2.5	48682	3	US-09-949-016-17236
24	41	2.5	48682	3	US-09-949-016-17237

C 25	41	2.5	48682	3	US-09-949-016-17238	Sequence 17238, A
C 26	41	2.5	48682	3	US-09-949-016-17239	Sequence 17239, A
C 27	41	2.5	51022	3	US-09-949-016-17135	Sequence 17135, A
C 28	41	2.5	51022	3	US-09-949-016-17136	Sequence 17136, A
C 29	41	2.5	51022	3	US-09-949-016-17137	Sequence 17137, A
C 30	41	2.5	51022	3	US-09-949-016-17138	Sequence 17138, A
C 31	41	2.5	86213	3	US-09-949-016-17240	Sequence 17240, A
C 32	41	2.5	86213	3	US-09-949-016-17241	Sequence 17241, A
C 33	41	2.5	86213	3	US-09-949-016-17242	Sequence 17242, A
C 34	41	2.5	86213	3	US-09-949-016-17243	Sequence 17243, A
C 35	40	2.5	1761	3	US-09-949-002-268	Sequence 268, App
C 36	40	2.5	1762	3	US-09-949-002-25	Sequence 25, Appl
C 37	40	2.5	1766	6	PCT-US93-00601-1	Sequence 1, Appl1
C 38	40	2.5	1766	6	PCT-US94-07107A-1	Sequence 1, Appl1
C 39	40	2.5	1767	3	US-08-083-945C-1	Sequence 1, Appl1
C 40	40	2.5	52202	3	US-09-949-016-17006	Sequence 17006, A
C 41	39.8	2.4	1141	3	US-09-806-708B-22	Sequence 22, Appl
C 42	39.6	2.4	168174	3	US-10-071-411A-63	Sequence 63, Appl
C 43	39.6	2.4	168273	3	US-10-071-411A-2	Sequence 2, Appl1
C 44	39.2	2.4	152331	3	US-09-128-155-16	Sequence 16, Appl
C 45	39.2	2.4	176373	3	US-09-128-155-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-09-492-029-2
; Sequence 2, Application US/09492029
; Patent No. 6929925
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Assays for Sensory Modulators Using a Sensory Cell
; TITLE OF INVENTION: Specific G-Protein Beta Subunit
; FILE REFERENCE: 02307E-092710US
; CURRENT APPLICATION NUMBER: US/09/492,029
; CURRENT FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: US 60/117,404
; PRIOR FILING DATE: 1999-01-27
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; TYPE: DNA
; LENGTH: 1520
; ORGANISM: Rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78)..(1097)
; OTHER INFORMATION: rat taste cell specific G-protein beta 3 subunit
; OTHER INFORMATION: (TC-Gbeta3)
US-09-492-029-2

Query Match	8.0%;	Score 130;	DB 3;	Length 1520;
Best Local Similarity	52.8%;	Pred. No. 8e-32;		
Matches	501;	Conservative	0;	Mismatches 390;
Indels	57;	Gaps	8;	
QY	383	GTACTCTTCAGGACACACCGGAAGTGTATTATTAGATTGGACACCGGAGGAACC	442	
Db	223	GGACGTTTAGGGACACCTGGCTAAGATCTATGCCATGCACTGGGCACCTCTTAGC	282	
QY	443	GGATTGTCACTGATCTCAAGATGGGAGATTATTCGTGGATGCTCTAACGAGTCAGA	502	
Db	283	TGCTAGTAGTGTCTCGCAGGATGGGAGCTGATGCTGTGGGACATTTACACCACCAATA	342	
QY	503	AAACTCATGTATTAACCTCCCTTGTGCATGGGTTATGACATGTGCTTCTTCTTCAATG	562	
Db	343	AGGTGATGCTATCCCGTCGCTTCCTCTCGGTGTCATGACCTGTGCTATGCACCATCAG	402	
QY	563	GTCACTGGTGTGCTGTGGTGGATTAGACAGTGTATGTTCTATCTTTAGCTTTAGCTCAA	622	

```
Db 403 GGAACCTTCGTGCATGTGGGGGCTAGATAACATGTGCTCAATCTACAGCCTCAAATC-- 460
Qy 623 CGGCGACAAGGATGGAATCTGTACCGTTTCAAGAATGCTCACTGGTCACAGGGGATATG 682
Db 461 -----CGTGAGGCAATGTCAAGGTTCAGCGCGGAACTCTCGGCTCACAGGTTATC 513
Qy 683 TTTCTGTCTGTCAGTATGTCCCAATAGAGGATGCCACCTTATCACAGGTTGAGTGATC 742
Db 514 TCTCTGTGTCCGCTTCCT-----GGATGCAACAACAATTTGACTAGCTCTGGGAGCA 567
Qy 743 AAACCTTGATCTTATGGGATGTAACCTACTGTGCTCAAACTTCTGTTTTTGGCGTGAT 802
Db 568 CCACTGTGCTTGTGGGCAATGTAGACGGGCGAGCAAGACAGTGTTG----- 618
Qy 803 TTCACTCTGGACATCTGCTGATGTACTAAGCGTCTCAATCAGTGGATCAAAACCCAAACT 862
Db 619 -----TGGGACACACTGTGTGACTCATGAGCTTGCTGT---GTCCCGAGACTACAAAC 669
Qy 863 GGTTTATATCTGGTTTCAAGCATTCACAGCAGCGTTGTGGACACTCGTGCTGCAAGCC 922
Db 670 TCTTCATCTCGGGAGCTTTGTGATGCCAGCGCCCAAGCTCTGGGATGTGAGGGAAGGCACT 729
Qy 923 GAGCAGTGTGCTACCTTTCAATGTACAGGGAGATGTTAATACGGTCAAGTTCTTTCCGG 982
Db 730 G-----TCGCCAGACTTCTACTGSCCAGAGTCHAGACATCATGCTATCTGTTTCTTCCCA 786
Qy 983 ATGGGTATAGATTTGGGACTGGATCAGAGATGGAACATGCAAGGCTGTATGACATAAGGA 1042
Db 787 ATGGGAGGCCATCTGCACTGGCTCAGATGATGCCCTCGCGCTCTTTGACCTGAGGG 846
Qy 1043 CTGGTCAACCACTCAGGCTCTATCAGCCACATGTTGATGTTGAGAGCGACCTGTCACT 1102
Db 847 CAGACCAGGAATGACAGCTTACTCCAC-----GAGACATCATCTGTGGCATCACT 900
Qy 1103 CCATTGCATTCCTGTGTGAGGAGACTTCTTTTCGCTGGCTATCGAGCAACAACACTT 1162
Db 901 CGTAGCCTTCTCACTCAGTGTGCTGCTCTTTGCTGGCTATGATGACTTCAAC---T 957
Qy 1163 GCTACGTTTGGGATACCTCTTTGGGAGAGTTGTATTTGGATTGGGATTTACAGCAGGAT 1222
Db 958 GCAATGTCTGGGACTCTGAAAGTGTGAGCGGTAGGGCTT-----CTTTCTG 1005
Qy 1223 CACACAGGAATAGATAGCTGTTTGGGGTTGTACAGAGATGGAAGTGCATGTGTACAG 1282
Db 1006 GCCATGACAACAGACTCAGTTGCTCGGGGTTCACAGCTGACGCGCATGGCTGTGGCCACTG 1065
Qy 1283 GAAAGTTGGGATCAAATCTAAAGATATGGCGTTTGGGAGGACACAGGA 1330
Db 1066 GATCTCGGACAGCTTCTCTCAAGATCTGGAATCTGAGGAGGCTGGAGGA 1113
```

RESULT 2

```
US-09-803-653-1
; Sequence 1, Application US/09803653
; Patent No. 6924100
; GENERAL INFORMATION:
; APPLICANT: Sifert, Winfried
; TITLE OF INVENTION: GENE ALTERATION IN THE GENE FOR THE GBETA3-SUBUNIT OF THE HUMAN G
; FILE REFERENCE: 741135-12
; CURRENT APPLICATION NUMBER: US/09/803,653
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: PCT/EP99/06534
; PRIOR FILING DATE: 1999-09-06
; PRIOR APPLICATION NUMBER: DE 198 41 299.1
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: DE 199 04 825.8
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: DE 199 12 049.8
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: DE 199 14 229.7
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: DE 199 19 989.2
; PRIOR FILING DATE: 1999-04-30
```

; PRIOR APPLICATION NUMBER: DE 199 23 539.2

; PRIOR FILING DATE: 1999-05-21

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 1

; LENGTH: 1523

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-803-653-1

Query Match 6.8%; Score 110.8; DB 3; Length 1523;

Best Local Similarity 51.3%; Pred. No. 1.9e-25;

Matches 495; Conservative 0; Mismatches 412; Indels 57; Gaps 8;

```
Qy 380 GTCGTACTCTTCAGGGACACACCGAAAGGTTTATTTCATTAGATTGGACACCGGAGGA 439
Db 149 GCGGACGTTAAGGGACACCTGCCAAGATTTCAGCCATGCATCTGGCCACTGATTCTA 208
Qy 440 ACCGATTTGTCAGTGCATCTCAAGATGGGAGATTAAATCGTGTGGAATGCTCTAACGATC 499
Db 209 AGCTGCTGTTAAGTCCCTCGCAAGATGGAAAGCTGATCGTGTGGACAGCTACACACCA 268
Qy 500 AGAAACTCATGCTATTAAACTCCCTTGTGTCATGGGTTATCACATGTGCTTCTCTCAA 559
Db 269 ACAAGTGCACGCCATCCACTCGGCTCCTCTGGGTCATGACCTGTGCTTATGCCCAT 328
Qy 560 ATGTCAGTCCGTTGCTGTGCTGATTTAGACAGTGTATGTTCTTATCTTTAGCCCTTAGCT 619
Db 329 CAGGNACTTTGTGGCATGTGGGGGCTGGACAACATGTGTTCCATCTACAACTCAAT 388
Qy 620 CAACGGCGGACAAAGATGGAATCTGACCGGTTTCAAGAAATGCTCCTCAGTGTGACAGGGAT 679
Db 389 C-----CCGTGAGGGCAATGTCAAGGTTCAGCGGGAGCTTTCTGCTCACACAGGT 439
Qy 680 ATGTTTCGTGCTGCTCAGTATGTCCTCCAAATGAGGATGCCACTTTATCACAGTTCAAGTG 739
Db 440 ATCTCTCTGCTGCGCTTCCT-----GGATGACAAACAATTTGTGACCAGCTCGGGGG 493
Qy 740 ATCAAACTTTGATCTTATGGGATGTAACTACTGCTCTCAAAACTTCTGTTTTTGGCGGTG 799
Db 494 ACACCACTGTGCTTGTGGACATTTGAGACTGGGACGACAGAACTGTATTG----- 547
Qy 800 AATTTCACTGTGGACATFACTCTGATGTACTAAGCGTCTCAATCAGTGGATCAAAACCAA 859
Db 548 -----TGGGACACACGGGTGACTGCATGAGCTGGCTGTGTCT---CTGACTTCA 595
Qy 860 ACTGTTTATATCTGTTTCATGCGATTTCCACAGACGGTTGTGGGACACTCGTGTCTCAA 919
Db 596 ATCTCTTCAITTTGGGGGCTGTGATGCCATGTGCCAAGCTCTGGGATGTGGAGAGGGGA 655
Qy 920 GCCGAGCAGTGCCTTACCTTTTCATGCTCACGAGGGAGATGTTAATACGGTCAAGTTCTTTC 979
Db 656 CCTG---CCGTGACACTTTCATGCGCACAGAGTGGACATCAACGGCATCTGTTTCTTCC 712
Qy 980 CGGATGGGTATAGATTTGGGACTCGATCAGACGATGGAACATGAGGCTGTATGACATAA 1039
Db 713 CCAATGGAGAGGCCATCTGACCGGCTCGGATGACGCTTCTCGCGCTGTTTGGACCTGC 772
Qy 1040 GGACTGGTCAACCAACTCCAGTCTCATGACCACATGTTGATGTTGAGAACGGACCTGTCA 1099
Db 773 GGGCAGACCAGAGCTGATCTGCTTCTCCC-----ACGAGAGCATCATCTCGCGCATCA 826
Qy 1100 CCTCATTTGCAATCTCTGTGTGTCAGGAGACTTCTTTTTCGCTGGCTATGCGAGCAACAACA 1159
Db 827 CGTCCGTGGCTTCTCCCTCAGTGGCGCCTACTATTTCGCTGGCTACGACGACTTCAAC- 885
Qy 1160 CTTGCTAGCTTTGGGATACCTCTTTGGGAGAGGTTGTATTGGATTGGGATTTACAGCAGG 1219
Db 886 --TGCAATGTCTGGGACTCCATGAAGTCTGAGCGTGTGGGCTCTCTC----- 932
Qy 1220 ATTCACACAGGAATAGAAATAGCTGTTTGGGGTTGTGAGGATGGAAGTGCATTGTGTA 1279
Db 933 -TGCCACGATACAGGGTGAAGTGCCTTGGGAGTTCACAGCTGACGGGATGCTGTGGCCA 991
```

QY 1280 CAGGAGTTGGGATCAAAATCTAAAGATATGGCGCTTTGGAGGACACAGGAGAGTGATTT 1339
 DB 992 CAGGTTCTTGGGACAGCTTCTCTCAAAATCTGGAACTGAGGAGCTGGAGAAAGGAAAGTG 1051
 QY 1340 GAAG 1343
 DB 1052 GAAG 1055

RESULT 3
 US-09-492-029-4
 ; Sequence 4, Application US/09492029
 ; Patent No. 6929925
 ; GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Adler, Jon Elliot
 ; APPLICANT: Lindemeier, Juergen
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Assays for Sensory Modulators Using a Sensory Cell
 ; TITLE OF INVENTION: Specific G-Protein Beta Subunit
 ; FILE REFERENCE: 02307E-09271005
 ; CURRENT APPLICATION NUMBER: US/09/492,029
 ; CURRENT FILING DATE: 2000-01-26
 ; PRIOR APPLICATION NUMBER: US 60/117,404
 ; PRIOR FILING DATE: 1999-01-27
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 1523
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (7)..(1029)
 ; OTHER INFORMATION: human taste cell specific G-protein beta 3 subunit
 US-09-492-029-4

Query Match 6.8%; Score 110.8; DB 3; Length 1523;
 Best Local Similarity 51.3%; Pred. No. 1.9e-25;
 Matches 495; Conservative 0; Mismatches 412; Indels 57; Gaps 8;
 QY 380 GTCTGACTCTTACGGACACACCGGAAAGTTTATTCATTAGATTGGACACCGGAGAGA 439
 DB 149 GCGGACGTTTAAAGGGACACCTGCGCAAGATTTACGCCATGCACTGGGCCACTGTCTA 208
 QY 440 ACCGATTGTCTAGTCATCTCAAGATGGGAGATTAATCGTGTGAATGCTCTAACGATC 499
 DB 209 AGCTGCTGTAAGTGCTTCCAAAGATGGGAGCTGATCGTGTGGACAGCTACACCA 268
 QY 500 AGAAACTCATGCTATTAAACTCCCTTGTGTCATGGGTTATGACATGCTTTCTCTCAA 559
 DB 269 ACAAGGTGACGCCATCCCACTGGCTCTCTGGTCAATGACCTGTGCTATGCCCAT 328
 QY 560 ATGCTGATCGGTTGGTGTGGTATGACAGTGTATGTTCTATCTTTAGCTTAGCT 619
 DB 329 CAGGAACTTTGTGGCATGTGGGGGCTGGCAACATGTGTTCCATCTACAACCTCAAT 388
 QY 620 CAACGGCGGACAGGATGGAATGTACCGGTTTCAAGATGCTCACTGGTTCACAGGGAT 679
 DB 389 C-----CCGTGAGGGCAATGTCAAGCTCAGCCGGAGCTTTCTGCTCACACAGTT 439
 QY 680 ATGTTTCTGCTGCTAGTATGTCCCAATAGGATGCCACCTTATCACCAGTTCAGGTG 739
 DB 440 ATCTCTCTGCTGCGCTTCTCT-----GGATGACAAATATTTGACCACTCGGGG 493
 QY 740 ATCAACTTGTATCTTATGGATGTAACTACTGCTCTCAAACTTCTGTTTGGGGGTG 799
 DB 494 ACACCACTGTGCTTTGTGGGCATTTGAGACTGGGACGAGAACTGTATTTG----- 547
 QY 800 AATTTCACTGCGACATCTGCTGATGATTAAGCGTCTCAATCAGTGGATCAACCCAA 859
 DB 548 -----TGGGACACACCGGGTGACTGCAATGAGCTTGGCTGTGCTCT---CCTGACTCA 595

QY 860 ACTGTTTATATCTGGTTTCATGCGATTCCACAGCACGGTTTGGGACACTGCTGCTGCAA 919
 DB 596 ATCTCTTATTTTCGGGGGCTTGTATGCGAGTCCAGTCCAGTCTGGGATGTGGAGAGGGGA 655
 QY 920 GCCGAGCAGTGGTACCTTTTCATGCTACAGGAGGAGATGTTAAATACCGTCAAGTTCTTTC 979
 DB 656 CTG---CCGTGAGACTTTTCTGCGCACGAGTCCGACATCAACGCCATCTGTTTCTTCC 712
 QY 980 CGGATGGGTATAGATTGGGACTGATCAGACGATGGAAATGCGAGGCTGTATGACATAA 1039
 DB 713 CCAATGGAGAGGCCATCTGCAAGGCTCGGATGAGCTTCTCTGCGCTTGTGTCACCTGC 772
 QY 1040 GGACTGGTCAACCACTCCAGTCTATCAGCCACATGTTGATGTTGAGAACGACCTGTCA 1099
 DB 773 GGGCAGACGAGAGCTGATCTGCTTCTCC-----ACGAGAGCATCATCTGCGGCATCA 826
 QY 1100 COTCCATTGCAATCTCTGTCAGGAGACTTCTTTTCGCTGGCTATGCGAGCAACAACA 1159
 DB 827 GTCTGCTGGCTTCTCTCCCTCAGTGGCGCCTACTATTCTGCTGCTACGAGACTTCAAC- 885
 QY 1160 CTTGCTACGTTTGGGATACCTCTTTGGGAGAGGTTGTTATTTGGATTGGGATTACAGCAGG 1219
 DB 886 --TGCATGCTGGGACTCCATGAAGTCTGAGCGTGTGGGCATCTCTC-----932
 QY 1220 ATTACACAGGATAGATTAAGCTGTTTGGGCTTGTGACGAGATGGAAGTCAATTGTGTA 1279
 DB 933 -TGGCCACGATTAACAGGGTGAGCTGCTGGGAGTCAACAGTCAAGGATGGCTGTGGCCA 991
 QY 1280 CAGGAAAGTTGGGATTTCAAAATCTAAAGATATGGGCTTTGGAGGACACAGGAGAGTATT 1339
 DB 992 CAGGTTCTTGGGACAGCTTCTCAAAATCTGGAATCTGAGAGGCTGGAGAAAGGAGTG 1051
 QY 1340 GAAG 1343
 DB 1052 GAAG 1055

RESULT 4
 US-09-949-016-1956
 ; Sequence 1956, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1956
 ; LENGTH: 3078
 ; TYPE: DNA
 ; ORGANISM: Human
 US-09-949-016-1956

Query Match 6.7%; Score 109.4; DB 3; Length 3078;
 Best Local Similarity 51.8%; Pred. No. 8.7e-25;
 Matches 481; Conservative 0; Mismatches 391; Indels 57; Gaps 8;
 QY 383 GTACTCTTACGGACACACCGGAAAGTTTATTCATTAGATTGGACACCGGAGAGAAC 442
 DB 425 GGACACTGGGGGGGACCTTGGCCAGATCTAGCCATGCACTGGGCAACAGACTCCAGGC 484
 QY 443 GGATTCTGATGCTATCTCAAGATGGGAGATTAATCGTGTGGAATGCTCTAAACGAGTCAG 502

```

Db 485 TTCTCGTCACTGCTCGCCAGGATGTTAACTTATCATCTGGGACAGCTACACCAACA 544
Qy 503 AAACATCATGCTATTAAACCTCCCTTGTGCAATGGGTTATGACATGTTCTTCTCCAAATG 562
Db 545 AGGTCCAGCCATCCCTCTCGCTCTCTCGGTGATGACCTGTCATATGCCCTTCTG 604
Qy 563 GTCACTCGGTTGGTGTGTGATAGACAGTGTATGTTCTATCTTTAGCTTACGTCAA 622
Db 605 GGAACATATGTGCGCTCGGTCGCTGGATTAACATTTGCTCCATTTACATCTGAAAC-- 662
Qy 623 CGGCGGACAAGATGGAACCTGTAACCGGTTTCAAGATGCTCACTGTCACAGGGATG 682
Db 663 -----TCGTGAGGGAAACGTGCGGTGAGTCGTGAGCTGGCAGACACAGGTTACC 715
Qy 683 TTTCGTGCTGTCAGTATGTCCTCAATATGAGATGCCACCTTATCACCAAGTTCAGGTGATC 742
Db 716 TGTCTGCTGCTCGGATTCCT-----GGATGACAAATCAGATCTGTCAGGCTGAGGACA 769
Qy 743 AAACATGATCTTATGAGGATGTAACATCTGCTGCTCAAACTTCTGTTTGGCGGTGAAT 802
Db 770 CCACGTGTCCTGCTGCTGAGATCGAGACCGGCGCAGACGACGACGCTTT----- 819
Qy 803 TTCAGTCTGGACATCTGCTGATGTAAGGCTCTCAATCAGTGTGATCAAAACCAACT 862
Db 820 -----ACCGGACACATGAGATGTCATGAGCTTTCTCT---TGCTCTGACACAGAC 871
Qy 863 GGTATATCTGTTTCACTGATTCACAGACGCTTGTGGACACATCTGCTGCTCAAGCC 922
Db 872 TGTCTGCTCTGCTGCTGATGCTTCTGATCTTCTGATCTGATCTGATCTGATCTGAT 928
Qy 923 GAGCAGTGGTACTCTTCAATGCTGAGGAGATGTTAATACCGTCAAGTCTTCTTCGG 982
Db 929 TGTGCGCGGACACCTTCACTGCGCCACGAGTCTGACATCAATGCCATTTGCTTCTTCCAA 988
Qy 983 ATGGGTATAGATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1042
Db 989 ATGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1048
Qy 1043 CTGGTCACCAACTCCAGGCTTATACGCCACATGCTGATGCTGATGATGATGATGATGAT 1102
Db 1049 CTGACGAGGCTCATGATTTACTCC-----ATGACAACTATCTGCGGGATCACT 1102
Qy 1103 CAATTGATCTCTGTCAGGAGATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1162
Db 1103 CTGCTCTCTCTTCCAGAGCGGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1159
Qy 1163 GCTAGCTTGGGATACCTCTTGGGAGAGGTTGATTTGGATTTGGATTTACAGCAGGAT 1222
Db 1160 GCAAGCTTGGGATGACCTCAAGCGGACCGCGGCGAGGTGCTTT-----GGCTG 1207
Qy 1223 CACACAGGATAGAATAAGCTTTTGGGTTTGTGACGAGATGGAAGTGCATTTGTGACAG 1282
Db 1208 GGCATGACAAACCGCTGAGCTGCTGGGCTGACCTGACGATGCGATGCTGCTGGGACAG 1267
Qy 1283 GAAGTGGGATCAAACTTAAGATATGG 1311
Db 1268 GGTCTGGGATAGCTTCTCTCAAGATCTGG 1296

```

RESULT 5
 US-09-023-655-1294
 ; Sequence 1294, Application US/09023655
 ; Patent No. 6607879
 ; GENERAL INFORMATION:
 ; APPLICANT: Cocks, Benjamin G.
 ; APPLICANT: Susan G. Stuart
 ; APPLICANT: Jeffrey J. Seilhamer
 ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
 ; NUMBER OF SEQUENCES: 1508
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
 ; STREET: 3174 PORTER DRIVE

```

; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1294:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3088 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 931667
; US-09-023-655-1294

```

```

Query Match      6.7%; Score 109.4; DB 3; Length 3088;
Best Local Similarity 51.8%; Pred. No. 8.8e-25;
Matches 481; Conservative 0; Mismatches 391; Indels 57; Gaps 8;

Qy 383 GTACTCTTCAGGACACACCGGAAGGTTTATTCATTAGATTGGACACCGGAGGAGAAC 442
Db 426 GGACACTCGGGGGACCTGGCCAGATCTAGCCATGCACTGGGGCAACAGCTCCAGGC 485
Qy 443 GGATTGTCACTGATCTCAAGATGGGAGATTAATCGTGTGAATGCTCTAACGAGTCAGA 502
Db 486 TTCGTCAGTGCCTCGCAGGATGTAACCTTATCATCTGGACAGCTACACCAACA 545
Qy 503 AAACATCATGCTATTAACTCCCTTGTGATGGGTTATGACATGCTGCTTCTTCTCAAAATG 562
Db 546 AGGTCCACGCCATCCCTCTGCGCTCTCTGGGTGATGACCTGATGATGATGATGATGATG 605
Qy 563 GTCAGTCGTTGCGTGTGGTGTAGACAGTGTATGTTCTATCTTTAGCCCTTAGCTCAA 622
Db 606 GGAATATGTGGCTGCGGTGGCTGGATAACATTTGCTTCCATTTACATCTGAAAC-- 663
Qy 623 CGGCGGACAAGGATGGAACCTGTAACCGGTTTCAAGAACTGCTCACTGCTGTCAGGGGATG 682
Db 664 -----TCGTGAGGGAAACGTGCGGTGAGCTGCTGAGCTGGCAGGACACACAGGTTACC 716
Qy 683 TTTCGTGCTGTCAGTATGTCCTCAAAATGAGGATGCCACCTTATCACCAAGTTCAGGTGATC 742
Db 717 TGTCTGCTGCGGATTCCT-----GGATGACAAATCAGATCTGTCAGGCTGAGGACA 770
Qy 743 AAACATGATCTTATGAGGATGTAACCTGCTCTCAAACTTCTGTTTGGCGGTGAAT 802
Db 771 CCAGTGTGCCCTGTGGGACATCGAGACCGGCGGACGAGACCAACGCTTT----- 820
Qy 803 TTCAGTCTGGACATCTGCTGATGATGTAACGCTCTCAATCAGTGGATCAAAACCAAACT 862
Db 821 -----ACCGGACACATCGAGATGTCATGAGCTTTCTCT---TGCTCTGACACAGAC 872
Qy 863 GGTATATCTGTTTCAATGCGATTCACAGACCGGTTGTGGGACATCTGCTGCTGCAAGCC 922

```


Query Match	5.1%	Score 83.8	DB 3	Length 1394
Best Local Similarity	58.0%	Pred. No. 1.5e-16		
Matches 148	Conservative 0	Mismatches 107	Indels 0	Gaps 0
Qy	380	GTCTACTCTTCAGGGACACACCGGAAGGTTTATTCAATTAGATTGGACACCGGAGAGGA	439	
Db	143	GCGGAGCTTAAAGGGGACACTGGCCAAAGATTTAAGCCATGCATCTGGGCCCATGATTCTTA	202	
Qy	440	ACCGGATTGTCAGTGCATCTCAAGATGGGAGATTAAATCGTGTGGAATGCTCTAACGAGTTC	499	
Db	203	AGCTGTGTTAAGTGCCTCGCAAGATGGAAGCTGATCGTGTGGGACAGCTACACCA	262	
Qy	500	AGAAAATCATGCTATTAAACTCCCTGTGTGCATGGGTTATACATGTGCTTCTCTCCAA	559	
Db	263	ACAAAGTGTCAGCCATCCCATCTGGGCTCCTCTGGGTCATGACCTGTGCCATATGCCCAT	322	
Qy	560	ATGTCAGTCGGTTGCGTGTGGATTAGACAGTGTATGTTCTATCTTTAGCCTTAGCT	619	
Db	323	CAGGGAACTTTGTGGCATGTGGGGGCTGGACACACATGTGTTCCATCTACAACCTCAAT	382	
Qy	620	CAACGGCGGACAAGG	634	

Db 383 CCCGTGAGGGCAATG 397

RESULT 11

US-09-032-372-6
; Sequence 6, Application US/09032372
; Patent No. 6008337
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Yue, Henry
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: CELL CYCLE RELATED PROTEINS
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/032,372
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0478 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2119 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNORAB01
; CLONE: 78191

US-09-032-372-6

Query Match 4.4%; Score 71.4; DB 3; Length 2119;
Best Local Similarity 57.3%; Pred. No. 2.7e-12;
Matches 129; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 383 GTACTCTTCAGGGACACACCGGAAAGGTTTATTTCATTAGATTGGACACCGGAGAGGACC 442
Db 362 GGACCTCAAGGCCCGGAAAGTCCTGTGCATGAGCTGGTGCAAGATAAGAGA 421
QY 443 GGATTGTGAGTGCACTTCAAGATGGGAGATTAAATCGTGTGGAATGCTCTAACGAGTCAGA 502
Db 422 GGATCGTGAGTCGTGCACGATGGGAGGTGATCGTGTGGGATTCCTTCACCAACAACA 481
QY 503 AAACATCATGCTATTAACCTCCCTTGTGATGGGTTATGACATGCTTTCTCTCAAAATG 562
Db 482 AGGAGCACCGGTCAACCATGCCCTGCACGTGGGTGATGCGATGCTTTATGCCCATCGG 541
QY 563 GTCAGTCGGTTGCGTGTGGTATGACAGTGTATGTTCTATCT 607
Db 542 GATGTGCCATTGCTTGTGGTGGTTTGGATAAAGTGTTCGTGT 586

RESULT 12

US-09-016-434-1017
; Sequence 1017, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1017:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: CERVNOT01
; CLONE: 930839

US-09-016-434-1017

Query Match 3.7%; Score 61; DB 3; Length 659;
Best Local Similarity 53.0%; Pred. No. 3.3e-09;
Matches 124; Conservative 0; Mismatches 110; Indels 0; Gaps 0;

QY 383 GTACTCTTCAGGGACACACCGGAAAGGTTTATTTCATTAGATTGGACACCGGAGAGAAC 442
Db 312 GGACATCTCGGGGGCACCTGGCCAAAGATCTNCGCCNTGCACTGGGGCACAGACTCCAGGC 371
QY 443 GGATTGTGAGTGCACTTCAAGATGGGAGATTAAATCGTGTGGAATGCTCTAACGAGTCAGA 502
Db 372 TTCTCGTCAGTGCCCTCGCAGGATGGTAAACTTATCNTCTGGGACAGCTACGCCACCNACA 431
QY 503 AAACATCATGCTATTAAACTCCCTTGTGCATGGGTTATGACATGTGCTTTCTCTCCAAATG 562
Db 432 AGGTCCTCGGCATCCCTCTGGCTCTCTCGGGTCATGACCTGTGCATNTGCCCTTCTG 491
QY 563 GTCAGTCGGTTGCGTGTGGTGGATTAGACAGTGTATGTTCTATCTTTAGCCTTA 616
Db 492 GGAACATATGTGGCTGCGGTGGCCTGGATAACATTTGCTCCATTTACAATCTGA 545

RESULT 13

US-09-270-767-53/c

Search completed: December 28, 2005, 19:32:47
Job time : 324 secs

THIS PAGE BLANK (U8PT0)